Previously uncharacterized rectangular bacterial structures in the dolphin mouth

Natasha K. Dudek[1,2,11], Jesus G. Galaz-Montoya[3], Handuo Shi[3,4], Megan Mayer[5,12], Cristina Danita[3], Arianna I. Celis[4], Tobias Viehboeck[6,7], Gong-Her Wu[3], Barry Behr[8], Silvia Bulgheresi[6], Kerwyn Casey Huang[3,4,9], Wah Chiu[3,4,5], David A. Relman[1,4,9,10,13,*]

[1]Department of Medicine, Stanford University School of Medicine, Stanford, CA 94305, USA

[2]Department of Ecology and Evolutionary Biology, University of California, Santa Cruz, Santa Cruz, CA 95064, USA

[3]Department of Bioengineering, Stanford University, Stanford CA 94305, USA

[4]Department of Microbiology and Immunology, Stanford University School of Medicine, Stanford, CA 94305, USA

[5]Division of CryoEM and Bioimaging, SSRL, SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA

[6]Department of Functional and Evolutionary Ecology, University of Vienna, Vienna, Austria

[7]Center for Microbiology and Environmental Systems Science, University of Vienna, Vienna, Austria

[8]Department of Obstetrics and Gynecology, Stanford University School of Medicine, Stanford, CA 94305, USA

[9]Chan Zuckerberg Biohub, San Francisco, CA 94158, USA

[10]Infectious Diseases Section, Veterans Affairs Palo Alto Health Care System, Palo Alto, CA 94304, USA

[11]Present address: Quantori, Cambridge, MA 02142, USA

[12]Present address: Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115, USA

[13]Lead Contact

*Correspondence: relman@stanford.edu

ABSTRACT

Much remains to be explored regarding the diversity of uncultured, host-associated microbes. Here, we report the discovery of unusual rectangular bacterial structures (RBSs) in the mouths of bottlenose dolphins. DNA staining revealed multiple paired bands within RBSs that suggested cells dividing along the longitudinal axis. Cryogenic transmission electron microscopy and tomography revealed parallel membrane-bound segments, suspected to be cells, encapsulated by an S-layer-like periodic surface covering. RBSs displayed novel pilus-like appendages with bundles of threads splayed at the tips. Multiple lines of evidence suggested that RBSs are bacterial and distinct from the Neisseriaceae genera *Simonsiella* and *Conchiformibius*, with which they share similar morphology and division patterning, including genomic DNA sequencing of micromanipulated RBSs, 16S rRNA gene sequencing, and fluorescence *in situ* hybridization. Our findings highlight the diversity of novel microbial forms and lifestyles that await discovery and characterization using tools complementary to genomics such as microscopy.

KEYWORDS

dolphin, microbiota, bacterial morphology, rectangular morphology, pili, S-layer, microscopy, cryogenic transmission electron microscopy (cryoEM), cryogenic tomography (cryoET), metagenomics

INTRODUCTION

The earliest descriptions of the microbial world centered around the morphology and motility patterns of 'animalcules'[1]. In the centuries since Leeuwenhoek's revolutionary discovery, a vast diversity of microbial forms have been described, ranging from star-shaped bacteria in the *Stella* genus[2,3] to the multicellular fruiting bodies characteristic of Myxobacteria[4,5]. Morphology is a biologically important characteristic, often highly conserved and molded over time by selective pressures resulting from an organism's lifestyle and environmental context[6]. Indeed, cell morphology plays an important role in motility, nutrient acquisition, cell division, and interactions with other cells, including symbioses with hosts, all of which are strong determinants of survival[7]. As such, morphological and structural studies offer an appealing route by which to glean insight into microbial life forms and the mechanisms by which species function and affect their environments. Moreover, characterizing the structures and functions of the diverse range of microbes in uncharted branches of the tree of life provides an opportunity to broaden our understanding of evolution and may result in myriad applications in biotechnology and medicine, exemplified by the development of optogenetics[8] and CRISPR-based gene editing[9].

Genomics serves as a powerful lens through which to describe the microbial world. In recent years, metagenomic and single-cell genomic analyses have substantially increased the number of known microbial phylum-level lineages by a factor of nearly four in the bacterial domain[10-13]. Sequencing the genomes of newly discovered organisms has led to the discovery of new functional systems, types of protein variants, and lifestyles[11,14-17], illustrating the correlation between phylogenetic diversity and functional potential[18]. However, the applicability of such approaches is mostly limited to

proteins and regulatory systems homologous to those of well-characterized organisms; the prediction of phenotypes and functions that are truly novel and/or whose genetic basis is unknown generally requires complementary knowledge. Given the recalcitrance of most microbial species on Earth to laboratory culturing (in 2016, 72% of approximately phylum-level lineages lacked any cultured representative)[19], methods that do not require cultivated isolates, such as microscopy, offer an appealing, complementary route by which to study novel morphological and functional properties of uncultured lineages. For example, recent advances in cryogenic electron microscopy (cryoEM) and tomography (cryoET) have allowed three dimensional (3D) imaging of intact bacterial cells at nanometer resolution[20], leading to important advances in the discovery and characterization of new microbial structures[21,22]. At present, our knowledge of cell biology has been largely limited to observation of and experimentation on bacteria that can be cultured, and thus there exists a severe bias in our understanding toward organisms conducive to growth in laboratory conditions. The use of microscopy and non-sequencing-based techniques for studying "the uncultured majority" will be essential for characterizing the full diversity of bacterial lifestyles that have evolved.

Despite the diversity of microbial cell shapes, rectangular structures are a rarity, with a poorly understood genetic basis. Such structures are of two types: individual cells that are rectangular, and cell aggregates that form rectangles. To the best of our knowledge, the discovery of non-eukaryotic rectangular cells has thus far been restricted to the family Halobacteriaceae, which consists of halophilic Archaea. Known rectangular cells from this family include *Haloquadratum walsbyi*[23], *Haloarcula quadrata*[24], and members of the pleomorphic genus *Natronrubrum*[25]. FtsZ-based fission has recently been

observed in the cube-shaped nematode symbiont *Candidatus* Thiosymbion cuboideus[26] and additional rectangular cells believed to be bacterial or archaeal have been discovered in high salinity environments but not taxonomically identified[27]. Among eukaryotic microorganisms, diatoms can have a rectangular appearance when visualized in two dimensions, although these cells are cylindrical rather than rectangular prisms[28]. A variety of bacteria form rectangular cell clusters, such as sheets of coccoid bacteria (for example, *Thiopedia rosea* and the genus *Merismopedia*[29]), and cuboidal structures of coccoid bacteria (for example, the genera *Sarcina* and *Eucapsis*[29]), and rectangular trichomes formed by disc-shaped bacteria (for example, *Oscillatoria limosa* and other cyanobacteria[29]).

Also rare in the microbial world are cells that diverge from the typical pattern of cell division along a transverse axis. Two spectacular examples are *Candidatus* Thiosymbion oneisti and *Candidatus* Thiosymbion hypermnestrae[32-34], which have been exclusively found on the surface of marine nematodes. This division patterning is thought to preserve attachment to the host[32,33]. Similarly, members of the Neisseriaceae genera *Alysiella*, *Simonsiella*, and *Conchiformibius* divide longitudinally, which is thought to help with adherence to human epithelial cells in the oral cavity[35]. These taxa are further striking in that they can be regarded as multicellular bacteria. Additional examples of bacteria that undergo longitudinal division include *Spiroplasma poulsonii*[36] and *Candidatus* Kentron, a clade of symbionts hosted by the marine ciliate *Kentrophoros*[37,38]. Such insight into the reproductive methods of diverse bacteria is essential for building a comprehensive understanding of cell biology.

Previous studies using 16S rRNA gene amplicon sequencing and genome-resolved metagenomics found that the mouths of bottlenose dolphins (*Tursiops truncatus*) host a rich collection of novel microbes and functional potential[16,39]. This diversity includes representatives from poorly characterized bacterial and archaeal phyla[19], some of which lack cultured representatives altogether and whose biology is therefore particularly poorly understood. To investigate the morphology and potential lifestyles of these novel microbial lineages, we surveyed dolphin oral microbial communities using various microscopy techniques. We discovered unusual rectangular bacterial structures (RBSs) in dolphin oral samples and characterized their cellular dimensions and DNA patterning using phase-contrast and fluorescence microscopy. Regular bands of DNA suggested that the units are sheets of individual cells, each of which is encapsulated by an inner and outer membrane, while fluorescence *in situ* hybridization (FISH) experiments and metagenomic sequencing strongly suggested that they are bacterial, and potential members of the Betaproteobacteria. Using cryogenic transmission electron microscopy (cryoTEM) and tomography (cryoET), we characterized the envelope structure of RBSs and discovered previously unobserved surface features such as heterogeneous bundles of appendages that protrude from the ends of RBSs and splay out at the tips. These findings highlight the power of high-resolution microscopy for exploring the nature of uncultivated microbes.

RESULTS

Rectangular bacterial structures in the dolphin oral cavity are Gram-negative and contain multiple bands of DNA

Oral swab samples were collected from the mouths of eight bottlenose dolphins (*Tursiops truncatus*) under the purview of the U.S. Navy Marine Mammal Program in San Diego Bay, California, USA, during three distinct intervals in 2012, 2018, and 2022 (**Methods**). Rectangular bacterial structures (RBSs) were readily apparent in phase contrast microscopy images (**Fig. 1a-e**). The RBSs resembled rectangular prisms; they were not cylindrical (**Supplementary Movie 1**). They exhibited Gram-negative characteristics following Gram staining (**Fig. 1f**). Attempts at membrane staining using FM4-64 were unsuccessful, as the FM4-64 dye did not stain any part of the RBSs. RBSs contained multiple parallel bands of fluorescence with DAPI staining (**Fig. 1b-e**). In some RBSs, the neighboring DNA band pairs appeared "H"-like (**Fig. 1g**, white arrow), suggesting two rod-shaped cells in the process of division along a longitudinal axis with the DNA bands undergoing segregation[26]. RBSs clustered into two morphotypes based on the dimensions of the rectangular units (**Fig. 1h**). Using materials from a single dolphin oral swab sample, we quantified the length and width of 23 RBSs. Morphotype A exhibited a median length of 3.95 ± 2.89 µm median absolute deviation (MAD) and median width of 5.08 ± 0.10 µm MAD (n=15 RBSs). Morphotype B had a median length of 3.08 ± 0.93 µm MAD and a median width of 2.21 ± 0.56 µm MAD (n=8 RBSs). The dimensions for both length and width were significantly different between the two morphotypes (for length, p=0.02; for width, $p=10^{-10}$; two-sided Welch's t-test). The different morphotypes may represent different cell types or taxonomic groups (for example, strains or species), cells in different stages of development, or cells with altered shape in response to environmental conditions.

We assessed the prevalence of RBSs of each morphotype in a set of 73 oral swab samples collected from eight dolphins (**Supplementary Table 1**), using a high-

throughput, automated phase contrast microscopy workflow that collected imaging data for 226 fields of view per sample[40]. RBSs of morphotype A, henceforth referred to as RBS-As, were detected in 39/73 samples from 7/8 dolphins (note that the number of samples per dolphin is not constant and that different oral locations were swabbed in different years). RBSs of morphotype B, henceforth referred to as RBS-Bs, were detected in 42/73 samples from 6/8 dolphins. Of 25 of these 73 samples that were collected from distinct oral sites, RBSs were detected in palatal (RBS-A = 5/5; RBS-B = 4/5) and gingival samples (RBS-A = 12/15; RBS-B = 14/15), but less frequently in buccal samples (RBS-A = 0/5; RBS-B = 2/5). Thus, we infer that the RBSs are stable colonizers in the dolphin oral cavity and have preferred colonization locations.

In this study, we focused on RBS-A, as this morphotype had a higher abundance in the dolphin oral samples and is more morphologically distinct from other, known microbes.

RBSs are likely bacterial and not Neisseriaceae

Given the intriguing morphology of RBSs, we next sought to determine their taxonomic affiliation. RBSs morphologically resemble members of the genera *Simonsiella* and *Conchiformibius* (Neisseriaceae family), which are rod-shaped oral commensals in mammals[30]. In a re-analysis of the Sanger clone library and 454 pyrosequencing data from a previous 16S rRNA gene amplicon survey of gingival swab samples from 38 dolphins from the same population[39], no *S. muelleri* (the sole species of the genus *Simonsiella*) or *Conchiformibius* amplicons were detected in any of these samples, although other members of the Neisseriaceae family were detected in these dolphin oral samples. A putative *S. muelleri* sequence type was detected in the Sanger library from the mouth and gastric fluid of one sea lion examined in the same amplicon study[39]

(NCBI accession number JQ205404.1); this partial 16S rRNA gene sequence has 94.6% identity over 99% query coverage to the partial *S. muelleri* ATCC 29453 16S rRNA gene sequence (NCBI accession number NR_025144.1). The former study detected five other Neisseriaceae sequence types in the Sanger library from sea lions (mouth and stomach), water, and fish species fed to marine mammals[39]. Meanwhile, there were four Neisseriaceae sequence types recovered in the 454 pyrosequencing survey, from dolphins (stomach, respiratory system), sea lions (mouth, stomach), fish, and seawater. These positive identifications indicate that *S. muelleri* and Neisseriaceae DNA could be extracted successfully in the previous study, although they were not detected in any dolphin oral samples[39].

We then performed 16S rRNA gene amplicon sequencing on 54 dolphin oral samples (Methods), resulting in the detection of 1,116,394 amplified sequence variants (ASVs) from 5,339,751 sequence reads. Of these 54 samples, 48 were screened in a high-throughput manner for RBSs via phase contrast microscopy (**Methods**). RBS-As and RBS-Bs were each detected in 22/48 samples (45.8%), but not the same 22 samples. One or the other morphotype was detected in 30/48 samples (62.5%). Rarefaction curves suggested that the depth of sequencing was sufficient for results to be near saturation for ASV richness (**Figure 2a)**. No Neisseriaceae amplicons were detected in these samples (**Figure 2b**), despite our ability to recover *S. muelleri* sequences from previously negative oral swab samples after deliberately spiking aliquots of these samples with *S. muelleri*. ASVs common to the RBS-A and RSB-B positive samples can be seen in **Supplementary Tables 2 and 3**, respectively.

The marine origin and rectangular nature of the RBSs also gave rise to speculation that they may be marine diatoms (for example, *Skeletonema costatum*), as cylindrical marine diatoms can appear rectangular in two dimensions. Thus, we next evaluated a potential eukaryotic origin of RBSs. We performed fluorescence *in situ* hybridization (FISH) using labeled eukaryotic (Euk-1209) and bacterial (Eub-338) probes, the latter of which is known to hybridize with both bacteria and archaea. As controls, we cultured and included *S*. *costatum* and the bacterium *Escherichia coli*. The Eub-338 probe hybridized to both *E. coli* and the RBSs, while the Euk-1209 eukaryotic probe hybridized to *S. costatum* cells alone (**Supplementary Fig. 1**), indicating that RBSs are not eukaryotic and thus not diatoms.

With no further *a priori* hypotheses as to the specific nature of the RBSs, we pursued a variety of general approaches to shed light on their identity. First, we cultured oral samples under aerobic and anaerobic conditions in three different media used to grow diverse bacteria (**Methods**), hoping to enrich for RBSs. Unfortunately, RBSs were not visible upon inspection of cultures under a microscope, indicating that the growth requirements for RBSs are distinct from those of typical bacteria isolated from mammalian microbiotas.

We further explored the potential for culturing RBSs directly on solid surfaces. Among the swab samples from 2022, we stored five in 20% glycerol immediately after collection to maximize the chances of maintaining RBS viability. Single-cell microscopy identified two glycerol stocks containing samples with numerous RBSs. These stocks were inoculated on agar plates containing either BSTSY-FBS or BHI-blood medium (**Methods**). BSTSY-FBS plates enable growth of *S. muelleri* and were incubated

aerobically at 37 °C. BHI-blood plates are typically used to grow a variety of bacterial commensals from mammals; these plates were incubated anaerobically at 37 °C. After 3 weeks of extended incubation, no colonies were visible on the BSTSY-FBS plates, while a control sample of *S. muelleri* formed large colonies after 1-2 days. The BHI-blood plates collectively contained ~300 colonies, of which we examined 288 using high-throughput single-cell microscopy[40]. None of the colonies contained cells with morphology similar to RBSs. Taken together, while disappointing from the point of view of enabling genomics approaches, our inability to successfully culture RBSs provides further support that they are not in the *Simonsiella* genus, which is readily cultivatible using such approaches.

Our next strategy employed mini-metagenomics, an approach in which a small number of cells are subsampled from a complex community and their DNA is amplified using multiple displacement amplification. Notably, this approach largely avoids preconceived biases about possible identity, since metagenomic analyses should reveal DNA from any cell from any domain of life, assuming successful cell lysis. We used three techniques to capture RBS-As for genomic sequencing: laser capture microdissection, microfluidics, and cell micromanipulation. Due to their large size compared with other bacteria, low density and the propensity of RBS-As to stick to other cells and to abiotic surfaces, only micromanipulation led to successful RBS-A capture (**Supplementary Fig. 1**). In addition to four collection tubes each containing ~1-3 RBS-As, four negative-control tubes of sample fluid were collected with the micropipette without any cells visible at the resolution of our microscope. Cell-free DNA and small non-target cells were also likely collected along with the RBS-As, given the frequent close proximity of RBS-As to other cells. DNA from RBS-A-positive and RBS-negative samples was

amplified using multiple displacement amplification (MDA), co-assembled, and sorted into 18 genome bins (**Fig. 4**; **Supplementary Fig. 3; Supplementary Tables 4, 5**; **Methods**). Notably, no *S. muelleri*, Neisseriaceae, or diatom genomes were recovered, although positive controls for these taxa were not included in the experimental design. The eukaryotic bins matched the human genome or the fungal class Malasseziomycetes, members of which are known commensals of human skin[41], and hence likely represent contaminants. No archaeal bins were recovered.

As half of the candidate bacterial bins recovered from the mini-metagenomics experiment were members of the Proteobacteria, we next performed FISH experiments using a set of class-specific probes targeting Alpha-, Beta-, and Gammaproteobacteria. The RBSs exhibited positive binding only to Betaproteobacteria probes (**Fig. 3**).

We synthesized insights into RBS-A taxonomic identity from each experimental line of evidence (**Fig. 4**). While no conclusive insights can be drawn about RBS-A identity, FISH results suggested a putative affiliation with the Betaproteobacteria. We note that a Betaproteobacteria partial genome from the family Alcaligenaceae was recovered from the mini-metagenomics experiment; the affiliation of this bin (number 16) with Alcaligenaceae was confirmed via phylogenetic analysis of the 16S rRNA and ribosomal protein S3 genes (**Supplementary Fig. 4, 5, Supplementary Methods**). An ASV consistent with the 16S rRNA gene of that bin is present in 11/13 of samples in which RBS-As were visually confirmed to be present and for which amplicon sequencing data are available.

The evolutionary path towards multicellularity and longitudinal division is poorly understood. Recent efforts to identify the genetic basis of these bacterial characteristics in the Neisseriaceae found that the acquisition of the amidase-encoding gene *amiC2*, along with modifications to key regulatory genes (e.g., *mreB*, *ftsA*), likely plays an important role. We searched the bins recovered from the mini-metagenomics experiment for putative AmiC2 proteins (those encoding Amidase_3, PF01520); candidates were identified in bins 4, 7, 9, 10, 11, 12, and 16 (the last of which is affiliated with Alcaligenaceae). Upon confirmation of the identity of the RBSs, future work should examine the significance of *amiC2* in conferring their unusual morphology.

CryoEM and cryoET reveal nanoscale surface and internal structures of RBS-As

To gain high-resolution structural insight into RBS-As, we imaged dolphin oral samples containing high densities of RBS-As using cryogenic transmission electron microscopy (cryoTEM). Low-magnification cryoTEM images revealed that each RBS-A consists of seemingly paired segments organized in parallel (**Fig. 5a,b**). These segments were oriented similar to the DAPI-stained bands seen in fluorescence microscopy images (**Fig. 1c,e**). Some groups of segments appeared to be in the act of separating from other groups, although our static data cannot definitively say whether these observations were reflective of cell division. Segments were surrounded by a dense, membrane-like layer under a low-density layer (**Fig. 6a,b (right) and Fig. 7a; Supplementary Movies 2, 3**).

We hypothesized that RBSs are most likely aggregates of cells, with each DNA-containing segment corresponding to an individual cell. The following observations support this hypothesis: 1) each individual segment appeared to be surrounded by an

inner and outer membrane, reminiscent of the plasma and outer membrane seen in other Gram-negative bacteria (**Fig. 5a, 6a,b (right) and Fig. 7a**); 2) segments are arranged in the same geometry (**Fig. 5a,b**) as the DAPI-stained bands and FISH probe-hybridized bands (**Fig. 1c,e,g**); 3) appendages protruded from the surface of individual segments (**Fig. 5c,d**); 4) RBS-As often consisted of variable numbers of segments that appeared to be separating from one another (**Fig. 1d,e and Fig. 6b**), suggesting that the rectangular structures do not reflect individual cells; 5) neighboring segments appeared H-shaped (**Fig. 1g**), reminiscent of nucleoids segregating in a bacterial cell undergoing longitudinal division; and 6) while a recent report describes the first discovery of a bacterium in which DNA is stored in a membrane-bound compartment[42], the number of known bacterial species in which DNA is physically segregated from the cytoplasm, let alone organelle-bound, is extremely low. By contrast, multicellularity has been documented in diverse bacterial species (reviewed in ref. [43]).

Dark, spherical structures were visible in the body of RBSs in cryoTEM images. In one tomogram, two dense spheroidal objects (**Fig. 5c**) were prominently visible and measured 192 nm × 200 nm × 192 nm (volume 3.1×10^{7} nm^3) and 215 nm × 220 nm × 220 nm (volume 4.4×10^{7} nm^3). Vesicle-like structures were also apparent. Notably, a surface covering with a periodicity of ~7-9 nm encapsulated the RBS-As (mean 7.83±0.86 nm SD) (**Fig. 6 and Fig. 7; Supplementary Movies 2, 3**). This feature was measured from the high-resolution 2D micrographs of two RBS-As from a single dolphin oral sample by generating line-density plots of segments (2 from each image) and manually measuring the distance between intensity peaks for 6-7 peaks.

To obtain more detailed three-dimensional reconstructions of RBS-A features, we conducted cryoET experiments. Tilt-series acquisitions were limited to the RBS-A periphery since the thickness of the RBS-A bodies occluded the electron beam almost completely at high tilt angles. The thickness at the RBS-A periphery (<1 µm from the edge) ranged from ~323 nm to ~751 nm, with an average value of ~509±132.4 nm SD (*n*=15), (above the ~500 nm threshold commonly regarded as the upper limit for productive cryoET imaging[44]). Notably, tilting increases the mean free path (MFP) of electrons in the electron beam, that is, the apparent ice thickness through which they travel. This increase is in inverse proportion to the cosine of the tilt angle, resulting in up to 3-fold the thickness at 70° compared to that at 0°. However, the fact that the electron beam is not completely occluded in two-dimensional images without tilting, as seen in our cryoEM images, suggests that the thickness of the RBS-A body is under the ~1,650 nm thickness penetrance limit for the 300 kV acceleration voltage of the microscope we used to image the RBS-As, although useful information was deemed to be completely lost at thicknesses of 650 nm or greater in benchmarks using microcrystal electron diffraction[45].

Appendages that resembled pili[46,47] protruded from RBS-As; these appendages often consisted of hair-like structures that formed bundles and splayed out at the tips, sometimes intertwining and/or crossing over one another (**Fig. 6b and Fig. 7b,c**). The bundles of appendages were structurally heterogeneous, with variable lengths, bundle widths, and numbers of tips. Notably, in examining the various features within the tomograms, we did not observe any membrane-bound organelles reminiscent of a nucleus, in line with a non-eukaryotic identity.

For both the appendages and periodic surface covering, subtomogram averaging[48] did not yield consistent maps, likely due to the thickness of the RBS-A periphery (often >500-600 nm thick), low signal-to-noise ratio of the tomograms, and limiting characteristics of the features in question, such as the variable curvature of the regions with stretches of continuous periodic surface covering. Successful subtomogram averaging typically relies on averaging identical structures, for example, repeated copies of a macromolecular complex, such as ribosomes in the same functional state. One can compensate for structural variability in the form of conformational and compositional heterogeneity with large datasets composed of thousands of subtomograms in combination with advanced classification methods. However, in our datasets, both the pilus-like appendages and the S-layer-like surface feature were structurally heterogenous and sparsely distributed in the RBS-As (e.g., the S-layer-like surface feature is not continuous along the entire membrane of the RBS-As, and in the stretches where it is, it exhibits differential curvature) and thus were observed only in a fraction of our tomograms.

To address sample thickness, we used cryogenic focused ion beam (FIB) scanning electron microscopy (SEM)[49] to mill thinner lamellae of RBS samples[50]. However, we could not locate any RBS-As under the ice, for two possible reasons: 1) with an inferred thickness between ~0.6 - ~1.7 µm, RBS-As may not form “mounds” under the ice that are protuberant enough to suggest where to mill; and 2) other larger cells in the samples (such as dolphin epithelial cells) formed more prominent mounds that obscured the RBS-As. Indeed, none of the lamella we produced from candidate locations contained RBS-As. The data from this experiment suggest that cryo-correlative light and electron microscopy (cryoCLEM) will be necessary in future studies to locate candidate

regions in cryoEM grids with vitrified RBS-As from which to produce thin lamellae using cryoFIB-SEM. Nonetheless, the cell surface exhibited features similar to those we observed on the RBS-A periphery, namely pili and S-layers. We suggest that the pilus-like appendages and the S-layer-like periodic surface covering of RBS-As merit future investigation.

DISCUSSION

Here, we used optical microscopy, cryoTEM, and cryoET to search for novel morphological diversity within the microbiota of dolphin oral samples. Morphological diversity was predicted based on previous findings of novel phylogenetic diversity and functional potential in the dolphin mouth via sequencing-based studies[16,39]. Interestingly, we discovered unusual RBSs. We infer that both RBS morphotypes are indigenous to the dolphin mouth given that they were consistently present in this environment: RBS-As and RBS-Bs were identified in 39/73 and 42/73 samples, respectively, that were surveyed using high-throughput microscopy imaging, including in 7/8 and 6/8 dolphins included in this study, and were present in samples collected during intervals ten years apart. Previous studies have found that the microbiota of marine mammals is distinct from that of seawater (even that of skin, which is constantly in direct contact with seawater)[39,51], and thus it is unlikely that RBSs are simply contaminants from seawater.

The taxonomic identification of specific cell morphotypes from complex communities can be extremely difficult, to the point that it often remains unresolved[27,52,53]. Results collected here strongly suggest that the RBS-As are not affiliated with the multicellular longitudinally dividing Neisseriaceae genus *Conchiformibius*, whose members also form rectangularly shaped clusters of rod-like cells[54]. First, the marker gene amplicon

sequencing workflow employed here did not detect any Neisseriaceae amplicons in 54 dolphin oral samples that underwent deep 16S rRNA gene amplicon sequencing, although this workflow did detect *S. muelleri* after cells from this taxon were intentionally spiked into aliquots of dolphin oral samples. Second, attempts to culture RBS-As using techniques that were employed successfully in our laboratory to culture *S. muelleri* failed, suggesting that the RBS-As have different physiological requirements for growth than those required by *S. muelleri*. Third, no Neisseriaceae genomes were recovered from the mini-metagenomics experiment. Fourth, visual comparisons of RBS-A images presented here with the TEM and fluorescence microscopy images of multicellular longitudinally dividing *Neisseriaceae* (*Alysiella*, *Simonsiella*, and *Conchiformibius*) in ref. [54] suggest that they are different taxa. For example, RBSs appear to contain segments (cells) that are embedded in a matrix-like material and fully encapsulated by an S-layer-like structure, whereas multicellular longitudinally dividing Neisseriaceae belonging to the same filament only appeared to share their outer membrane[54].

FISH experiments indicated that the RBS-As are bacterial and likely members of the Betaproteobacteria. A Betaproteobacteria genome was recovered from the mini-metagenomics experiment, from the family Alcaligenaceae. An ASV matching the 16S rRNA gene of this bin was detected in 11/13 samples that underwent amplicon sequencing and were visually confirmed to contain RBS-As via phase contrast microscopy. Taken together, the RBS-As may be members of the Alcaligenaceae, although the finding that this ASV was absent in 2/13 samples confirmed to contain RBS-As via microscopy challenges this hypothesis. One possibility is that the RBS-As were present in sufficiently low relative abundance in those two samples as to not be detected, despite deep amplicon sequencing. Another possibility is that RBS-As are not

this Alcaligenaceae taxon, and rather the Alcaligenaceae taxon is a ubiquitous member of the dolphin oral microbiota and therefore showed up frequently in our sequencing-based analyses. In such a case, it would suggest that either the RBS-As are Betaproteobacteria that did not lyse in the sequencing-based experiments or that the FISH results were a false positive, despite the stringent and controlled conditions under which the experiment was performed.

Obtaining a species-level identification for RBSs via sequencing-based methods will be extremely challenging for numerous reasons, such as the frequent close proximity of RBSs with other small cells that were likely mixed with RBSs during micromanipulation or recalcitrance to laboratory lysis. Importantly, we are hesitant to exclude candidate identities based on not being present in the mini-metagenomics experiment in all positive RBS samples, since technical limitations could have resulted in false negatives. For example, a thick cell wall or obstruction preventing reagents from reaching the RBS by the micropipette needle could have interfered with lysis of the cell membrane and impeded DNA extraction. Conversely, a positive result in a negative control may have arisen due to non-specific read mapping, contamination of genome bins with material from true contaminants, or cell-free DNA. Culturing-based approaches may ultimately be the most promising route forward for identifying RBSs, although many combinations of parameters will likely need to be tested to find satisfactory conditions for RBS growth. Regardless of the taxonomic identity of the RBSs, the novel structural features that have evolved in these microorganisms are intriguing and highlight the discovery potential for further study.

The paired nature of segments in RBSs can likely be ascribed to their longitudinal mode of binary fission, as seen in the Neisseriaceae genera *Alysiella*, *Simonsiella*, and *Conchiformibius*, as well as *S. poulsonii, Ca. T. oneisti, and Ca. T. hypermnestrae*[32,33,36,55]. In the Neisseriaceae member *S. muelleri*, sheets are thought to help cells remain physically anchored in the oral cavity when rapidly shedding epithelial cells slough off[35]. We hypothesize that the same may be true for RBSs, which inhabit a similar environment and whose morphology may have undergone convergent evolution due to similar evolutionary pressures. Longitudinal binary fission may be an even more general characteristic that is selected in response to the need to form a secure attachment to a substrate. The segments at the ends of RBSs are often shorter than those closer to the center, suggesting that there may be a mechanism by which the growth of segments is determined by their spatial positioning within an RBS. The RBSs present another case example for future evolutionary studies focused on understanding the drivers and genetic basis of bacterial multicellularity and longitudinal division.

CryoTEM images suggested that RBS-As are encapsulated by a periodic surface covering, which may be an S-layer or a new crystalline structure. S-layers are self-assembling, crystalline arrays of single proteins or glycoproteins that coat the exterior of some bacteria and archaea[56,57]. While their exact function varies widely across microorganisms and is often unknown, S-layers are hypothesized to confer beneficial functions given their high metabolic cost (the S-layer comprises up to ~20% of the total protein synthesized by cells), their ubiquity across microbes, and their multiple evolutionary origins[56,57]. If segments in RBS-As correspond to individual cells, production of the periodic surface covering may represent cooperation between cells within RBS-As. Cooperative synthesis of a single, shared periodic surface covering by

multiple cells could have evolved since close kin (other cells in an RBS-A) have limited dispersal ability and are therefore situated in close physical proximity. RBS-As would benefit from cooperative production of a single periodic surface covering around a population of cells rather than around each individual segment by reducing the surface area required to cover all cells, and such an advantage could even have contributed to selection for aggregation. An additional and not mutually exclusive possibility is that the periodic surface covering may help to maintain the ultrastructure of segments within an RBS-A, similar to archaea such as *Thermoproteus tenax*[58].

One of the most striking features of RBS-As is their pilus-like appendages. At present, there are five characterized classes of pili in Gram-negative bacteria (chaperone-usher, curli fibers, F-type, type IV, and type V) and two general types of pili in Gram-positive bacteria (short, thin rods and longer, flexible, hair-like filaments)[46,47]; other pilus-like appendages have been documented, such as hami in archaea[21]. To the best of our knowledge, characterized bacterial pili all consist of single appendages that exist as independent units. By contrast, the pilus-like appendages that protrude from RBS-A segments exhibit an unusual architecture involving heterogeneous bundles of filaments that often splay out at the tips. These observations raise the question of whether the RBS-A appendages represent a novel type of assembly of pilin subunits, or are a completely distinct class of appendages.

Extensive investigation of hundreds of cryoEM images and dozens of cryoET tomograms enabled visualization of the structure of many RBS-A features at nanometer resolution. Future studies of RBS-As may benefit from imaging with phase-plate optics that dramatically increase image contrast[59] following specimen preparation methods

that thin cells into lamellae by focused ion beam (FIB) milling coupled with scanning electron microscopy (SEM) at cryogenic temperatures[49,50]. Our data suggest that cryoCLEM will be needed to enable production of thin lamellae for RBS-As since the RBS-A cell body seems to be thicker than the limit allowed for cryoET experiments, and yet too thin for RBS-As to be readily found by cryoSEM prior to cryoFIB milling without fluorescent labels. Successful cryoCLEM+cryoFIB-SEM experiments followed by cryoET could enable more comprehensive analyses of the community of RBS-As and their cell body beyond the thin periphery as well as visualization of subcellular components of interest at higher resolution via subtomogram averaging.

The vast majority of microorganisms on Earth lack isolated representatives[19]. Sequencing-based analyses have proved invaluable in exploring and describing said diversity, yet cannot be used to explore all aspects of the biology of microorganisms. Notable blind spots in our understanding of uncultured organisms include the unique genes and corresponding structural and functional features that have evolved within these lineages. While the use of advanced imaging techniques to visualize microbes can provide insight into the biology of uncultured lineages, a shift toward a more multifaceted approach drawing on many disciplines and techniques will be required to create a comprehensive view of this biological dark matter[60,61].

Acknowledgments

We thank members of the Relman lab for insightful discussion, Katharine Ng and Brian Yu for assistance in attempts to capture RBS-As using laser capture microdissection and microfluidics, respectively, Melissa Clark for assistance with Gram staining, Michael Schmid for analysis of the periodic surface covering on RBS-As, and especially Celeste

Parry and colleagues at the U.S. Navy Marine Mammal Program in San Diego, CA for collecting dolphin oral samples. This work was supported by NIH grant P41GM103832 (W.C.), a James S. McDonnell Postdoctoral Fellowship (H.S.), the Austrian Science Fund (T.V., S.B.), and the Thomas C. and Joan M. Merigan Endowment at Stanford University (D.A.R.). K.C.H. and D.A.R. are Chan Zuckerberg Biohub Investigators.

Author contributions

Conceptualization: N.K.D, K.C.H., D.A.R.

Methodology and investigation: N.K.D., J.G.G.M., H.S., M.M., C.D., G.H.W, T.V., S.B., B.B., K.C.H, W.C., D.A.R.

Formal analysis: N.K.D, J.G.G.M., H.S., C.D.

Visualization: N.K.D, J.G.G.M, H.S., C.D.

Writing - original draft: The manuscript was mainly written by N.K.D. and J.G.G.M. based on N.K.D.'s original draft, with revisions by all other authors.

Writing - review and editing: all authors.

Supervision: K.C.H, W.C., D.A.R.

Competing interests

The authors declare no competing interests.

METHODS

Data and code availability

Sequencing data for this project is available through NCBI BioProject PRJNA174530. Raw reads for the amplicon survey were deposited to SRA and are associated with BioSamples SAMNXXXXXXXX-SAMNXXXXXXXX. Raw reads for the single cell genomics experiment were also deposited to SRA; the captured RBSs and negative controls are physically derived from a single oral swab represented by BioSample SAMN19012476, while the reads each from the eight experimental replicates (four RBSs, four negative controls) are each individually associated with BioSamples SAMN19022663-SAMN190226970. The co-assembly of scaffolds ≥5kb in length from the single cell genomics experiment was deposited as a Whole Genome Shotgun project at DDBJ/ENA/GenBank under the accession JAHCSF000000000, following the removal of human-derived sequences. The version described in this paper is version JAHCSF010000000. Genome bins 2-5 and 7-18 from the single cell genomics experiment have been deposited as a Whole Genome Shotgun project at DDBJ/ENA/GenBank under accessions JAGYHI000000000-JAGYHX000000000. The versions described in this paper are versions JAGYHI010000000-JAGYHX010000000. Genome bin 1 (human) was not deposited. Scaffolds for genome bin6 (<100,000 nucleotides) were deposited as a non-genome GenBank submission under accession numbers MZ126582-MZ126593.

Experimental model and subject details. Oral swab samples were obtained from bottlenose dolphins (*Tursiops truncatus*) managed by the U.S. Navy Marine Mammal Program (MMP) Biosciences Division, Space and Naval Warfare Systems Center Pacific, San Diego, USA. The earliest sample containing RBSs was collected on April 1, 2012 and the latest on March 24, 2022. Swab samples were obtained using sterile foam Catch-All sample collection swabs (Epicenter, WI, Cat. #QEC091H). Samples collected

in 2012 were obtained by swabbing the left gingival sulcus. Samples collected in 2018 were obtained by swabbing the palate, the tongue, and the left gingival sulcus (all three surfaces for each swab). Samples collected in 2022 were obtained from the palate, buccal surface, or left gingival sulcus. Of the 2022 samples of the gingival sulcus, 5 were stored in 20% glycerol. All other swab samples were dry frozen. The swabbing protocol adhered to the guidelines described in the CRC Handbook of Marine Mammal Medicine.

The MMP is accredited by the Association for Assessment and Accreditation of Laboratory Animal Care (AAALAC) International and adheres to the national standards of the United States Public Health Service Policy on the Humane Care and Use of Laboratory Animals and the Animal Welfare Act. As required by the U.S. Department of Defense, the MMP's animal care and use program is routinely reviewed by an Institutional Animal Care and Use Committee (IACUC) and by the U.S. Navy Bureau of Medicine and Surgery. The animal use and care protocol for MMP dolphins in support of this study was approved by the MMP's IACUC and the Navy's Bureau of Medicine and Surgery (IACUC #92-2010, BUMED NRD-681).

Microscopy sample preparation. To separate cells from swabs, swabs were immersed in 1X PBS (~50-100 µL, depending on cell density) in microcentrifuge tubes. Tubes were vortexed vigorously for ~10 s and lightly centrifuged to remove liquid from tube caps. The resulting solution was used for microscopy.

Light microscopy. Approximately 1 µL of cell solution in PBS was spotted onto an agarose pad (1% agarose in PBS) and imaged with an Eclipse Ti-E inverted microscope

with a 100X (NA: 1.4) objective (Nikon, Tokyo, Japan). To determine DNA localization, cells were stained with DAPI at a final concentration of 0.5 µg/mL for 5 min prior to imaging using emission/excitation spectra of 340/488 nm. High-throughput, automated imaging of dolphin oral samples via phase contrast microscopy was achieved using the Strain Library Imaging Protocol (2016 implementation)[40] to capture 226 fields of view for each of the samples collected in 2018, and 100 fields of view for those collected in 2022.

Gram and FM4-64 staining. Gram staining was performed using a Gram Staining Kit (Sigma Aldrich, cat. #77730-1KT-F) following manufacturer's protocol. Cells were imaged under a bright field microscopy with 100X objective (Nikon). FM4-64 dye (ThermoFisher Scientific, cat. # T13320) was applied directly to dolphin oral swab samples following manufacturer's protocol. The FM4-64 dye did not stain any part of the RBSs.

RBS-A cryofixation and cryoEM/ET data acquisition. A solution of cells in PBS (2.5 µL) was applied to glow-discharged 200-mesh copper, holey-carbon Quantifoil grids (Quantifoil, Großlöbichau, Germany, Cat. #Q2100CR1) with holey carbon or gold GridFinder Quantifoil grids (Quantifoil, Großlöbichau, Germany, Cat. #LFH2100AR2), followed by application of 2 µL of 15-nm gold fiducial solution to both sides of each grid. Grids were blotted for 5 s and plunge-frozen in liquid ethane cooled by liquid nitrogen to approximately -195 °C using an EM GP Plunge Freezer (Leica, Wetzlar, Germany).

Samples were loaded into one of two microscopes: a Titan Krios G3 operated at 300 kV with an energy filter (20-eV slit width), or a Titan Krios G4 operated at 300 kV without an

energy filter. Both microscopes were equipped with a K2 Summit direct electron detection device (Gatan, Pleasanton, USA) used to record micrographs. Data were acquired semi-automatically in counting mode using SerialEM (v. 3.8)[62]. CryoEM/ET imaging parameters are provided in **Supplementary Table 6**.

CryoEM/ET data processing. Montages were blended and binned 4-fold or greater using IMOD's "blendmont" algorithm[63] and normalized, band-pass filtered, rotated, and cropped for display purposes using EMAN2 v. 2.39[64]. Fifteen out of sixteen tilt series were suitable for tomographic reconstruction in IMOD v. 4.12.9. Tilt series with sampling at 7.5 Å/pixel were down-sampled by 2-fold and those with sampling at 3.48 Å/pixel or 3.75 Å/pixel were down-sampled by 4-fold. Images with artifacts such as excessive charging, drifting, large ice contamination creeping in at high tilts, or excessive thickness at high tilt were excluded from 12 of the tilt series prior to manual gold-fiducial-based alignment; up to 13 images were removed out of the 41 images in the original raw tilt series.

Tomograms were reconstructed using standard weighted back-projection and a SIRT-like filter (mimicking 16 iterations) and were band-pass-filtered and further binned by 2-fold in most cases for feature annotation, segmentation, movie production, and other display purposes. Tomogram thickness was estimated by visually identifying the smallest and largest *z*-slices with visible RBS-A or ice contamination densities and converting the number of slices to nanometers. Subtomogram averaging was attempted using EMAN2 v. 2.39[65,66] for globular densities suspected to be ribosomes, matrix densities under the outer membrane, patches of the periodic surface covering, and regions of pilus-like appendages, but no interpretable structures with resolution better

than ~50 Å were obtained. The ranges of thickness and length for the pilus-like appendages were derived by visually scanning the slices in the tomograms for the thinnest individual filaments and thickest bundles perceptible to the naked eye, and measuring their dimensions in binned-by-4 tomographic slices using the measuring tape tool of EMAN2's `e2display.py`. The repeat distance of the periodic surface covering was measured manually in a similar fashion as the pilus-like appendages from tomographic slices, with ~10-20 measurements from each of three tomograms displaying at least small regions where the repeat was discernible. This quantification yielded a range between ~6 and 10 nm, suggesting that either the layer components are flexible or that the underlying structure can yield different apparent distances between its subunits depending on the angle at which it is sliced. Additionally, regions showing the pattern much more clearly in higher-magnification montage two-dimensional projection images were cropped out, rotated to lie in a horizontal plane, filtered, and masked to compute line-density profiles parallel to the outer membrane.

We initially carried out tomographic annotation of three features (periodic surface covering, lipid membranes, and pilus-like appendages) for three tomograms using EMAN2's semi-automated two-dimensional neural network-based pipeline[67] and performed manual clean-up of false positives in UCSF Chimera v. 1.16[68]. The output annotation probability maps from EMAN2 v. 2.39 were turned into segmentations by applying a visually determined threshold and multiplying the contrast-reversed tomograms by the thresholded annotation map. The segmentations were low-pass-filtered with EMAN2 v. 2.39 to smooth out noise. However, since the complexity of subcellular structures was not captured by the semi-automated annotations, we applied a similar process to generate segmentations of five features (pilus-like appendages,

periodic surface covering, outer membrane, matrix, and inner membranes) using manual annotations performed with IMOD v. 4.12.9, following a recent protocol that increases manual annotation efficiency[69]. Snapshots for **Fig. 6** displaying RBS-A features in color as well as **Supplementary Movies 2 and 3** showing segmentation results were produced with UCSF Chimera v. 1.16.

CryoFIB-SEM. We first identified a grid that contained RBS-As using light microscopy. Using an Aquilous Cryo-FIB (ThermoFisher Scientific, MA, USA), we created 5 lamellae using 30 kV, 30 pA current and 5 µs duration time. The samples were then loaded into a cryo-TEM for sample observation and data collection. We were unable to identify RBSs in the lamellae.

Fluorescence *in situ* hybridization. Cell cultures of axenic *Escherichia coli* MG1655, non-axenic *Skeletonema costatum* LB 2308 (UTEX Culture Collection of Algae at the University of Texas at Austin, Austin, TX, USA), *Caulobacter crescentus* CB15N, and *Simonsiella muelleri* ATCC29453 were prepared as controls. *E. coli* was cultured in LB broth and grown at 37 °C, *S. costatum* was cultured in Erdschreiber's Medium at 20 °C with a ~12 h light and ~12 h dark cycle, *C. crescentus* was cultured in PYE medium at 30 °C, and *S. muelleri* was cultured in BSTSY medium at 37 °C.

All FISH probes were ordered from Integrated DNA Technologies (Coralville, USA) with HPLC purification. Probe sequences and fluorescence labels are as follows: Euk-1209: 5'-GGGCATCACAGACCTG-/3Alx660/-3', Bact338: 5'-GCTGCCTCCCGTAGGAGT-/Alx488/-3', ALF968: 5'-/Cy5/-GGTAAGGTTCTGCGCGTT-3', BET42a: 5'-/Alx594/-

GCCTTCCCACTTCGTTT-3', GAM42a: 5'-/Alx488/-GCCTTCCCACATCGTTT-3', nonEUB: 5'-/Cy5/-ACTCCTACGGGAGGCAGC-3'.

Cells from controls and RBS-As were collected in microcentrifuge tubes. To ensure sufficient biomass from dolphin oral swabs, cells from four swabs were condensed into a single tube. The FISH protocol was adapted from ref. [70]. Cells were fixed in 1 mL of 3.7% formaldehyde solution (800 µL of DEPC-treated water, 100 µL of 10X PBS, 100 µL of 37% formaldehyde) for 30 min with gentle shaking at 700 rpm. Cells were then washed twice in 1 mL of 1X PBS, and permeabilized in a mixture of 300 µL of DEPC-treated water and 700 µL of 200-proof ethanol with gentle shaking at 700 rpm for 2 h. Probes were added to 50 µL of hybridization solution to a final concentration of 1 µM per probe set. For ALF968, BET42a, and GAM42a, the hybridization solution contained 55% formamide solution (4 mL of DEPC-water, 1 g of dextran sulfate, 4.85 mL of formamide, 1 mL of 2X SSC, brought to a total volume of 10 mL with DEPC-treated water); for other probes, the hybridization solution contained 40% formamide (5 mL of DEPC-water, 1 g of dextran sulfate, 3.53 mL of formamide, 1 mL of 2X SSC, brought to a total volume of 10 mL with DEPC-treated water). For ALF968, BET42a, and GAM42a, cells were incubated in 50 µL of hybridization buffer with probes at 46 °C for 1 h; for other probes, cells were incubated overnight in 50 µL of hybridization buffer with FISH probes at 30 °C. Cells were washed twice using a wash solution (2 mL of 20X SSC buffer, 7.06 mL of formamide, 10.94 mL of DEPC-treated water) and resuspended in 2X SSC buffer. One microliter of cells was mounted onto 1% agarose pads containing PBS and 5 µg/mL of DAPI for imaging. Imaging data were processed using FIJI v. 2.0.0[71].

16S rRNA amplicon sequencing and processing. Fifty-four dolphin oral samples were selected for 16S rRNA gene amplicon sequencing. Genomic DNA was extracted from dolphin oral samples and 32 negative (PBS) controls using the DNeasy UltraClean 96 Microbial Kit (Qiagen Cat. #10196-4) following manufacturer's instructions. The 16S V4 region was amplified using 515F and 806rB primers using Platinum™ II HotStart PCR Master Mix (ThermoFisher Cat. #14000013). The PCR products were pooled at equal volume and gel-purified. Final purification was performed using Macherey-Nagel NucleoSpin Gel and PCR Clean-up, Mini Kit (Fisher, Cat. #740609). Amplicons were sequenced on the Illumina MiSeq platform with 250-bp paired reads at the Stanford Chan Zuckerberg Biohub Facility**,** resulting in a median read depth of 92,077 reads (min: 50,772, max: 265,108).

Demultiplexing was performed using Bcl2Fastq v. 2 (Illumina, CA, USA). ASVs were inferred using DADA2 v. 1.16.0, following guidelines in the "Big Data Workflow" (https://benjjneb.github.io/dada2/bigdata_paired.html). Taxonomic affiliations were assigned using the SILVA 138 SSU database[72] as a reference. Forward and reverse reads were trimmed to 240 nt and 180 nt, respectively. This pipeline yielded a total of 1,116 taxa and 5,339,751 reads across the 54 samples. ASVs were analyzed using phyloseq v. 1.28.0[73].

Spike-in experiment. To confirm that our DNA extraction, amplification, sequencing, and bioinformatic analysis pipeline was able to identify members of the Neisseriaceae, we performed a spike-in experiment. We first selected a dolphin oral sample that contained RBS-As, as determined by phase-contrast microscopy. We then created two aliquots of this sample. To the first aliquot, we spiked in cells from an *S. muelleri* pure

culture at a 1:1 ratio. The other aliquot was left untouched. These two samples, along with an aliquot of the *S. muelleri* pure culture (positive control), underwent DNA extraction, PCR amplification, and sequencing protocol as described above. The three samples were sequenced in a single Illumina HiSeq run, which was not the same lane as the other amplicon samples from this study. Note that by PCR amplifying *S. muelleri* in the same lab environment as the negative control sample and then sequencing it on the same lane, some cross-contamination is to be expected. The first time the negative control sample was sequenced (in the absence of *S. muelleri* pure cultures in the laboratory), no Neisseriaceae amplicons were detected.

Mini-metagenomics. To obtain candidate identities for RBS-As, we employed a mini-metagenomics approach. To limit contamination by foreign DNA, reagents, tubes, and PBS were treated with 11.4 J/cm^2 of ultraviolet light following the guidelines in ref. [74]. RBS-As were visualized using an Olympus IX70 inverted microscope (Olympus, Waltham, USA) with a 40X objective and Hoffman modulation optics. An Eppendorf TransferMan micromanipulator (Eppendorf, Hamburg, Germany, Cat. #5193000020) with a SAS-10 microinjector was used to capture RBS-As with Polar Body Biopsy Micropipettes (30° angled, beveled, and polished with an inner diameter of 13-15 µm) (Cooper Surgical, Målov, Denmark, Cat. #MPB-BP-30). After an RBS-A or chain of RBS-As was acquired, the micropipette tip was transferred to a collection tube containing 1X PBS and crushed into the tube to ensure the RBS-A(s) was deposited in the tube; this precaution was adopted because RBS-As frequently stuck to the glass micropipette and could not be dislodged. No dolphin cells were captured, although cell-free DNA and small, non-target cells from the sample were likely acquired as contaminants along with RBS-As based on the propensity of the latter to attach to other

species (**Fig. 5e**). Four tubes of RBS-As were collected (sample names RBS1-4), along with four negative-control tubes (NEG1-4). Negative controls consisted of draws of PBS from the same sample that did not contain any visible cells and were otherwise treated identically to RBS-A-containing samples.

DNA from each tube was amplified via multiple displacement amplification (MDA) using the Repli-g single-cell kit (Qiagen, Hilden, Germany, Cat. #150343) according to the manufacturer's protocol. DNA was purified using a Zymo Clean and Concentrate Spin Column (Zymo Research Corporation, Irvine, USA, Cat. #D4013) and libraries were prepared using the Kapa Hyper Prep Kit (Kapa Biosystems, Wilmington, USA, Cat. #KK8504) at the W.M. Keck Center for Comparative Functional Genomics at the University of Illinois, Urbana-Champaign. The eight libraries were sequenced using the Illumina MiSeq 2x250 nt P2 V2 platform. RBS-A samples RBS1, RBS2, NEG1, and NEG2 were pooled and sequenced across a single lane that produced 11,371,243 read pairs, and samples RBS3, RBS4, NEG3, and NEG4 were pooled and sequenced across 1.5 lanes, collectively producing 19,615,690 read pairs. Sequencing adapters were computationally removed at the Keck Center.

Reads from all eight samples were co-assembled using SPAdes v. 3.11.1[75] with the single cell (–sc) and careful (–careful) modes specified. A total of 61,973,866 read pairs were used for assembly, resulting in 1,406 scaffolds ≥5 kbp long with a total length of 17,438,233 bp and an N50 of 14,592 for scaffolds ≥5 kbp long. Protein coding genes were identified using Prodigal v. 2.6.2[76]. Per scaffold average coverage was calculated by mapping reads per sample against the co-assembly using bowtie2 v. 2.2.4[77], using

the samtools v. 1.6 depth function[78] to calculate per-base read coverage, and a custom script to calculate average per-base read coverage per scaffold.

A search for the *amiC2* gene in bins recovered from the mini-metagenomics experiment was performed by querying Pfam alignment PF01520 against each genome's protein sequences, using HMMER suite v. 3.1b2[79].

To determine the taxonomic identity of sequenced cells, we employed a genome-resolved approach. Assignment of scaffolds to genome bins was performed using the tetranucleotide frequencies of all scaffolds ≥5 kbp long over windows of 5 kbp, as described in ref.[80]. Results were computed and visualized using the Databionics ESOM Tools software v. 1.1[81], leading to the reconstruction of 18 genome bins (**Supplementary Fig. 3**). To refine bins, we removed scaffolds for which <50% of keys were assigned to the bin. Scaffolds <5 kbp long were not binned. The completeness and contamination per bin were assessed using CheckM v. 1.0.7[82]. To evaluate how representative binning was of the genomes that were sequenced, we estimated the number of prokaryotic genomes expected to be recovered by searching the metagenome assembly for a set of 16 bacterial single copy genes (bSCGs) assumed to be present in every genome in a single copy[83]. Profile HMMs for relevant proteins were obtained from the Pfam database[84] and queried against our dataset using HMMER suite v. 3.1b2[79].The median number of each bSCG was 10, suggesting ~10 prokaryotic genomes were represented in our sequencing dataset. In the case of the Alcaligenaceae genome of interest, the 16S rRNA gene was manually extended from the end of a scaffold; using bowtie2 v. 2.2.4 we ensured that the reads supported the final sequence.

Taxonomic identification of bins posed a challenge since 16S/18S rRNA genes were not reliably amplified/sequenced/assembled, and genomes were partial with few phylogenetically informative bacterial single copy genes present in the dataset. Hence, we used BLAST v. 2.2.30[85] to query all protein coding genes from each genome against the NCBI non-redundant protein database using an e-value of 10^{-10} and taxonomic assignments were made based on the closest protein match. Genome bin taxonomic assignments were considered highly likely if ≥50% of the top BLAST hits originated from a single taxon and were considered plausible if <50% but ≥33% of the top BLAST hits originated from a single taxon.

There are numerous approaches by which one could assess whether a bin is "present" in a sample, each with largely arbitrary thresholds. We focused on the relative abundance of each bin per sample (**Supplementary Table 5**) as it accounts for the length of each bin and allows for comparisons between samples with different numbers of read pairs[16].

Attempt at culturing RBSs. Four dolphin oral samples confirmed to contain RBSs were selected for culturing efforts. For each sample, one milliliter of sterile PBS was added to a 1.5-mL Eppendorf tube containing the oral swab sample. For liquid culturing, 600 µL of each sample were used to inoculate 3 mL of BSTSY[86] (2.75% (w/v) Tryptic Soy Broth, 0.4% (w/v) yeast extract, 10% bovine serum), SHI[87], or mSHI media (SHI supplemented with 0.9 g/L NaCl, 2.5 g/L K_2PO_4, 0.84 g/L $NaHCO_3$, 0.17 g/L $CaCl_2$, 0.04 g/L $MgCl_2$•$6H_2O$, and 5 g/L dextrose). BSTSY was selected for its use in successfully culturing bacteria (*S. muelleri* specifically) from the oral cavities of various mammals[86].

SHI was selected because this medium was designed to sustain high diversity communities derived from the human oral microflora. mSHI (modified SHI) was included as a higher-salinity version of SHI in an attempt to further mimic the conditions that might be found in the oral cavity of dolphins. Inoculation was repeated under anaerobic conditions in an anaerobic chamber (COY Lab Products, Grass Lake, USA); note that all samples were unavoidably exposed to atmospheric oxygen prior to culturing. Cultures were incubated at 37 °C to mimic the body temperature of dolphins. No RBSs were detected in liquid media by visual screening under a microscope after ~24, ~48, ~72, and ~96 h. For solid-surface culturing, ~10^3-10^4 cells (verified by microscopy to contain RBS-As) were directly plated onto BSTSY or BHI-blood (BHI medium supplemented with 5% horse blood) agar plates, and incubated at 37 °C with or without oxygen, respectively. No colonies grew on the BSTSY plates after 3 weeks of incubation; the colonies grown on the BHI-blood plates were screened using microscopy and no RBSs were visible. By contrast, a control of *S. muelleri* streaked onto BSTSY plates developed visible colonies after 1-2 days of incubation with oxygen at 37 °C, and the colonies were verified under microscopy to consist of cells with the morphology expected of *S. muelleri*.

Quantification and statistical analysis. Detailed descriptions of the quantitative methods used in this paper can be found in the Results and Method Details sections. Briefly, these descriptions include the methods used for DNA extraction, sequencing, read quality filtering and analysis, and microscopy data acquisition.

FIGURES AND LEGENDS

Figure 1. Light microscopy reveals RBSs with multiple morphotypes and distinct DNA banding patterns. a) Phase-contrast images of RBSs (arrow indicates one example) on the surface of dolphin oral epithelial cells. See also **Supplementary Movie 1**. **b,c)** Some RBSs exhibited long bands of DAPI fluorescence. Phase-contrast image is shown in (**b**), with fluorescence overlay in cyan in (**c**). **d,e)** Other RBSs exhibited shorter DAPI bands. Phase-contrast image is shown in (**d**), with fluorescence overlay in cyan in (**e**). Dark spots (arrowheads) were organized in lines perpendicular to DAPI-stained bands. DAPI-stained bands appeared to be organized in pairs. **f)** Gram-stained RBSs display Gram-negative characteristics. Inset: Gram-stained *Bacillus subtilis* (*B.s.*, Gram-positive) and *Escherichia coli* (*E.c.*, Gram-negative). **g)** Neighboring DNA band pairs in an RBS form “H”-like shapes (white arrow), likely because the DNA bands are segregated nucleoids in a cell undergoing longitudinal division. **h)** The two RBS morphotypes have distinct distributions of length and width. The centers of the orange and blue crosses represent the means of RBS-As andRBS-Bs, respectively, while the lengths of the arms represent ±1 standard deviation.

Figure 2. 16S rRNA gene amplicon sequencing of dolphin oral samples indicates that RBSs are not affiliated with the Betaproteobacteria family Neisseriaceae. Dolphin oral samples (*n*=54) were subjected to deep 16S rRNA gene amplicon sequencing. **a)** Rarefaction curves for the 54 sequenced dolphin oral samples. **b)** For each family of Betaproteobacteria detected in the 54 dolphin oral samples, the relative abundance of member ASVs is plotted. Note that the genus *Simonsiella* is a member of the Betaproteobacteria family Neisseriaceae; no ASVs affiliated with this family were detected. Visual examination using phase-contrast microscopy (see **Methods**) revealed RBSs of either morphotype A or B in 39 of 73 samples screened (note that in total 73 samples were visually screened for RBSs; 54 were used for amplicon sequencing, while the remainder were used for other experiments) (**Supplementary Table 1**). **c)** Sequencing of an *S. muelleri* pure culture and a dolphin oral sample (confirmed to contain RBS-As) with *S. muelleri* spiked in resulted in the detection of a Neisseriaceae ASV. The same ASV was detected at a relative abundance of <0.1% in that same dolphin oral sample in its unaltered state (no *S. muelleri* added) when prepared and sequenced in parallel with the *S. muelleri* positive samples; the ASV was not detected in any dolphin oral sample prior to the introduction of *S. muelleri* into the lab environment.

Figure 3. Fluorescence *in situ* hybridization indicates that RBS-As are Betaproteobacteria. Probes for Proteobacteria groups Alphaproteobacteria (ALF968), Betaproteobacteria (BET42a), and Gammaproteobacteria (GAM42a) were assessed for their ability to hybridize to RBSs. Top row: phase-contrast images; middle three rows from top to bottom: fluorescence images for Alphaproteobacteria, Betaproteobacteria, and Gammaproteobacteria probes, respectively; bottom row: DAPI staining of presumed DNA. Arrows highlight relevant cells in samples. RBSs of the “A” morphotype (**Figure 1h**) hybridized with the Betaproteobacteria probe and exhibited minimal hybridization with the Alphaproteobacteria and Gammaproteobacteria probes.

Figure 4. Insights into RBS-A identity. Table shows the 18 bins recovered from MDA and sequencing of RBS samples collected via micromanipulator, along with the phylum from which they are inferred to derive. For lowest taxonomic identity achieved (or class, in the case of the polyphyletic Proteobacteria), an asterisk (*) denotes lower confidence in the assignment (Methods). The RBS-A panel shows relative abundances of bins in each of the four samples that contained RBS-As based on visualization, color-coded as follows: green: ≥5%, yellow: ≥1% and <5%, orange: >0% and <1%, red: 0%. The negative-control panel (Neg) presents the same information for each of the samples that did not appear to contain RBS-As. Criteria for gauging the likelihood of a bin deriving from the RBS-As are shown: (1) Was the bin ever present in negative controls (green = no, yellow = yes but never ≥1% relative abundance, orange = yes but never ≥5%, red = yes and at least once ≥5%)? (2) Of ASVs detected in >75% of samples that underwent 16S rRNA gene amplicon sequencing and were visually confirmed to contain >10 RBS-As (n=13), was an ASV of this taxonomic identity? Numbers in boxes indicate the number of samples in which this ASV was present. For recovered bins detected only to the level of Gammaproteobacteria, a large and diverse group, this criterion is marked in yellow and ASV counts are not shown. (4) Based on FISH results, which bins are supported as potential candidates for the RBS-As? A "†" denotes that members of the Gracilibacteria phylum are inferred not to be Gram-negative from genomic studies (although they are not necessarily Gram-positive)[88]. Highest likelihood candidates are green for all three criteria.

Figure 5. CryoTEM demonstrates that RBSs consist of multiple parallel, likely paired segments and are often near other cells. Band-pass filtered and denoised low magnification (**a,b**) images or (**c**) montage cryoTEM images of RBSs on an R2/2 holey carbon TEM grid. Higher-magnification images showing (**d**) pilus-like appendages and (**e**) proximal cells and interacting vesicles at the RBS periphery. In (**a**), pairs of segments are highlighted with alternating shades of blue, and sharp indentations between groups of segments are denoted with black arrows. Dense spheroidal objects were present inside RBSs (white arrow). In (**d**), segments are encapsulated by an inner membrane (IM) and outer membrane (OM), denoted by black arrows. **d,e**) Representative micrographs showing that RBSs were often in close physical proximity to other cells in the samples. In (**e**), an apparent small indentation (black arrow) in the RBS periodic surface covering overlaps with a non-RBS cell or vesicle. Small dark spots (15 nm) in (**c-e**) are gold fiducial particles used for tilt series alignment in cryoET experiments (black arrow).

Figure 6. CryoET reveals the three-dimensional architecture of RBS-A components. a,f) Examples of ~3-nm-thick slices at two different depths (left and center) from two representative RBS-A tomograms, and corresponding manual annotations of cellular features (right). The tomograms are thick (~500-600 nm) and as such, pilus-like appendages (yellow) were visible only at a certain depth within the tomographic volume density of the RBS-A. See also **Supplementary Movies 2 and 3.** Blue, inner membranes; purple, periodic surface covering; green, outer membrane; red, matrix.

Figure 7. RBS-A surface features include appendages in heterogeneous bundles that splay out at the tips and a periodic surface covering around the entire RBS-A. a) Single cryoTEM image. **b,c)** CryoET slices (~3-nm thick) of an RBS-A. Red and orange boxes in (**c**) are magnified views of bundles of appendages, and yellow arrows denote thin, single appendages. **d)** Representative 2D cryoTEM image at the edge of an RBS-A showing a periodic surface covering. **e)** Line density profiles along selected regions from the image in (**d**) show that the spacing of the repetitive features is ~7-9 nm along a direction parallel to the RBS-A membrane.

REFERENCES

1 Leeuwenhoek, A. V. Observations, communicated to the publisher by Mr. Antony van Leewenhoeck, in a dutch letter of the 9th Octob. 1676. here English'd: concerning little animals by him observed in rain-well-sea-and snow water; as also in water wherein pepper had lain infused. *Philosophical Transactions of the Royal Society of London* **12**, 821-831 (1677).

2 Nikitin, D., Vasilyeva, L. & Lokhmacheva, R. New and rare forms of soil microorganisms. *M.: Nauka* (1966).

3 Vasilyeva, L. Stella, a New Genus of Soil Prosthecobacteria, with Proposals for Stella humosa sp. nov. and Stella vacuolata sp. nov. *International Journal of Systematic and Evolutionary Microbiology* **35**, 518-521 (1985).

4 Dworkin, M. Introduction to the myxobacteria. *Prokaryotic Development*, 219-242 (1999).

5 Voelz, H. & Reichenbach, H. Fine structure of fruiting bodies of Stigmatella aurantiaca (Myxobacterales). *J Bacteriol* **99**, 856-866, doi:10.1128/JB.99.3.856-866.1969 (1969).

6 Young, K. D. The selective value of bacterial shape. *Microbiol Mol Biol Rev* **70**, 660-703, doi:10.1128/MMBR.00001-06 (2006).

7 Young, K. D. Bacterial morphology: why have different shapes? *Curr Opin Microbiol* **10**, 596-600, doi:10.1016/j.mib.2007.09.009 (2007).

8 Fenno, L., Yizhar, O. & Deisseroth, K. The development and application of optogenetics. *Annu Rev Neurosci* **34**, 389-412, doi:10.1146/annurev-neuro-061010-113817 (2011).

9 Ishino, Y., Krupovic, M. & Forterre, P. History of CRISPR-Cas from Encounter with a Mysterious Repeated Sequence to Genome Editing Technology. *J Bacteriol* **200**, doi:10.1128/JB.00580-17 (2018).

10 Anantharaman, K. *et al.* Thousands of microbial genomes shed light on interconnected biogeochemical processes in an aquifer system. *Nat Commun* **7**, 13219, doi:10.1038/ncomms13219 (2016).

11 Brown, C. T. *et al.* Unusual biology across a group comprising more than 15% of domain Bacteria. *Nature* **523**, 208-211, doi:10.1038/nature14486 (2015).

12 Castelle, C. J. *et al.* Genomic expansion of domain archaea highlights roles for organisms from new phyla in anaerobic carbon cycling. *Curr Biol* **25**, 690-701, doi:10.1016/j.cub.2015.01.014 (2015).

13 Rinke, C. *et al.* Insights into the phylogeny and coding potential of microbial dark matter. *Nature* **499**, 431-437, doi:10.1038/nature12352 (2013).

14 Burstein, D. *et al.* New CRISPR-Cas systems from uncultivated microbes. *Nature* **542**, 237-241, doi:10.1038/nature21059 (2017).

15 Donia, M. S. *et al.* A systematic analysis of biosynthetic gene clusters in the human microbiome reveals a common family of antibiotics. *Cell* **158**, 1402-1414, doi:10.1016/j.cell.2014.08.032 (2014).

16 Dudek, N. K. *et al.* Novel Microbial Diversity and Functional Potential in the Marine Mammal Oral Microbiome. *Curr Biol* **27**, 3752-3762 e3756, doi:10.1016/j.cub.2017.10.040 (2017).

17 Wrighton, K. C. *et al.* Fermentation, hydrogen, and sulfur metabolism in multiple uncultivated bacterial phyla. *Science* **337**, 1661-1665, doi:10.1126/science.1224041 (2012).

18 Wu, D. *et al.* A phylogeny-driven genomic encyclopaedia of Bacteria and Archaea. *Nature* **462**, 1056-1060, doi:10.1038/nature08656 (2009).

19 Hug, L. A. *et al.* A new view of the tree of life. *Nat Microbiol* **1**, 16048, doi:10.1038/nmicrobiol.2016.48 (2016).

20 Pilhofer, M., Ladinsky, M. S., McDowall, A. W. & Jensen, G. J. Bacterial TEM: new insights from cryo-microscopy. *Methods Cell Biol* **96**, 21-45, doi:10.1016/S0091-679X(10)96002-0 (2010).

21 Moissl, C., Rachel, R., Briegel, A., Engelhardt, H. & Huber, R. The unique structure of archaeal 'hami', highly complex cell appendages with nano-grappling hooks. *Mol Microbiol* **56**, 361-370, doi:10.1111/j.1365-2958.2005.04294.x (2005).

22 Tocheva, E. I., Li, Z. & Jensen, G. J. Electron cryotomography. *Cold Spring Harb Perspect Biol* **2**, a003442, doi:10.1101/cshperspect.a003442 (2010).

23 Walsby, A. E. A square bacterium. *Nature* **283**, 69-71 (1980).

24 Oren, A., Ventosa, A., Gutierrez, M. C. & Kamekura, M. Haloarcula quadrata sp. nov., a square, motile archaeon isolated from a brine pool in Sinai (Egypt). *Int J Syst Bacteriol* **49 Pt 3**, 1149-1155, doi:10.1099/00207713-49-3-1149 (1999).

25 Xu, Y., Zhou, P. & Tian, X. Characterization of two novel haloalkaliphilic archaea Natronorubrum bangense gen. nov., sp. nov. and Natronorubrum tibetense gen. nov., sp. nov. *Int J Syst Bacteriol* **49 Pt 1**, 261-266, doi:10.1099/00207713-49-1-261 (1999).

26 Weber, P. M. *et al.* FtsZ-mediated fission of a cuboid bacterial symbiont. *iScience* **25**, 103552, doi:10.1016/j.isci.2021.103552 (2022).

27 Alam, M., Claviez, M., Oesterhelt, D. & Kessel, M. Flagella and motility behaviour of square bacteria. *EMBO J* **3**, 2899-2903 (1984).

28 Horner, R. A. A taxonomic guide to some common marine phytoplankton. (2002).

29 Zinder, S. H. & Dworkin, M. Morphological and physiological diversity. *The prokaryotes. A handbook on the biology of bacteria, 3rd edn. Springer, New York*, 185-220 (2006).

30 Hedlund, B. P. & Kuhn, D. A. The genera Simonsiella and Alysiella. (2006).

31 Xie, C. H. & Yokota, A. Phylogenetic analysis of Alysiella and related genera of Neisseriaceae: proposal of Alysiella crassa comb. nov., Conchiformibium steedae gen. nov., comb. nov., Conchiformibium kuhniae sp. nov. and Bergeriella denitrificans gen. nov., comb. nov. *J Gen Appl Microbiol* **51**, 1-10, doi:10.2323/jgam.51.1 (2005).

32 Leisch, N. *et al.* Growth in width and FtsZ ring longitudinal positioning in a gammaproteobacterial symbiont. *Curr Biol* **22**, R831-832, doi:10.1016/j.cub.2012.08.033 (2012).

33 Leisch, N. *et al.* Asynchronous division by non-ring FtsZ in the gammaproteobacterial symbiont of Robbea hypermnestra. *Nat Microbiol* **2**, 16182, doi:10.1038/nmicrobiol.2016.182 (2016).

34 Pende, N. *et al.* Host-Polarized Cell Growth in Animal Symbionts. *Curr Biol* **28**, 1039-1051 e1035, doi:10.1016/j.cub.2018.02.028 (2018).

35 McCowan, R. P., Cheng, K. J. & Costerton, J. W. Colonization of a portion of the bovine tongue by unusual filamentous bacteria. *Appl Environ Microbiol* **37**, 1224-1229, doi:10.1128/AEM.37.6.1224-1229.1979 (1979).

36 Ramond, E., Maclachlan, C., Clerc-Rosset, S., Knott, G. W. & Lemaitre, B. Cell Division by Longitudinal Scission in the Insect Endosymbiont Spiroplasma poulsonii. *mBio* **7**, doi:10.1128/mBio.00881-16 (2016).

37 Foissner, W. Kentrophoros (Ciliophora, Karyorelictea) has oral vestiges: A reinvestigation of K-fistulosus (Faure-Fremiet, 1950) using protargol impregnation. *Arch Protistenkd* **146**, 165-179 (1995).

38 Seah, B. K. B. *et al.* Sulfur-Oxidizing Symbionts without Canonical Genes for Autotrophic CO2 Fixation. *mBio* **10**, doi:10.1128/mBio.01112-19 (2019).

39 Bik, E. M. *et al.* Marine mammals harbor unique microbiotas shaped by and yet distinct from the sea. *Nat Commun* **7**, 10516, doi:10.1038/ncomms10516 (2016).

40 Shi, H., Colavin, A., Lee, T. K. & Huang, K. C. Strain Library Imaging Protocol for high-throughput, automated single-cell microscopy of large bacterial collections arrayed on multiwell plates. *Nat Protoc* **12**, 429-438, doi:10.1038/nprot.2016.181 (2017).

41 Limon, J. J., Skalski, J. H. & Underhill, D. M. Commensal Fungi in Health and Disease. *Cell Host Microbe* **22**, 156-165, doi:10.1016/j.chom.2017.07.002 (2017).

42 Katayama, T. *et al.* Membrane-bounded nucleoid discovered in a cultivated bacterium of the candidate phylum ‘Atribacteria’. *bioRxiv*, 728279 (2019).

43 Lyons, N. A. & Kolter, R. On the evolution of bacterial multicellularity. *Current opinion in microbiology* **24**, 21-28 (2015).

44 Diebolder, C. A., Koster, A. J. & Koning, R. I. Pushing the resolution limits in cryo electron tomography of biological structures. *J Microsc* **248**, 1-5, doi:10.1111/j.1365-2818.2012.03627.x (2012).

45 Martynowycz, M. W., Clabbers, M. T. B., Unge, J., Hattne, J. & Gonen, T. Benchmarking the ideal sample thickness in cryo-EM. *Proc Natl Acad Sci U S A* **118**, doi:10.1073/pnas.2108884118 (2021).

46 Hospenthal, M. K., Costa, T. R. D. & Waksman, G. A comprehensive guide to pilus biogenesis in Gram-negative bacteria. *Nat Rev Microbiol* **15**, 365-379, doi:10.1038/nrmicro.2017.40 (2017).

47 Proft, T. & Baker, E. N. Pili in Gram-negative and Gram-positive bacteria - structure, assembly and their role in disease. *Cell Mol Life Sci* **66**, 613-635, doi:10.1007/s00018-008-8477-4 (2009).

48 Galaz-Montoya, J. G. & Ludtke, S. J. The advent of structural biology in situ by single particle cryo-electron tomography. *Biophys Rep* **3**, 17-35, doi:10.1007/s41048-017-0040-0 (2017).

49 Marko, M., Hsieh, C., Schalek, R., Frank, J. & Mannella, C. Focused-ion-beam thinning of frozen-hydrated biological specimens for cryo-electron microscopy. *Nat Methods* **4**, 215-217, doi:10.1038/nmeth1014 (2007).

50 Wu, G. H. *et al.* Multi-scale 3D Cryo-Correlative Microscopy for Vitrified Cells. *Structure* **28**, 1231-1237 e1233, doi:10.1016/j.str.2020.07.017 (2020).

51 Apprill, A., Mooney, T. A., Lyman, E., Stimpert, A. K. & Rappe, M. S. Humpback whales harbour a combination of specific and variable skin bacteria. *Environ Microbiol Rep* **3**, 223-232, doi:10.1111/j.1758-2229.2010.00213.x (2011).

52 Luef, B. *et al.* Diverse uncultivated ultra-small bacterial cells in groundwater. *Nat Commun* **6**, 6372, doi:10.1038/ncomms7372 (2015).

53 Wanger, G., Onstott, T. C. & Southam, G. Stars of the terrestrial deep subsurface: a novel 'star-shaped' bacterial morphotype from a South African platinum mine. *Geobiology* **6**, 325-330, doi:10.1111/j.1472-4669.2008.00163.x (2008).

54 Nyongesa, S. *et al.* Evolution of longitudinal division in multicellular bacteria of the Neisseriaceae family. *Nat Commun* **13**, 4853, doi:10.1038/s41467-022-32260-w (2022).

55 Steed, P. A. Simonsiellaceae fam.nov.with characterization of Simonsiella crassa and Alysiella filiformis. *J Gen Microbiol* **29**, 615-624, doi:10.1099/00221287-29-4-615 (1962).

56 Fagan, R. P. & Fairweather, N. F. Biogenesis and functions of bacterial S-layers. *Nat Rev Microbiol* **12**, 211-222, doi:10.1038/nrmicro3213 (2014).

57 Sleytr, U. B. *et al.* S-layers as a tool kit for nanobiotechnological applications. *FEMS Microbiol Lett* **267**, 131-144, doi:10.1111/j.1574-6968.2006.00573.x (2007).

58 Wildhaber, I. & Baumeister, W. The cell envelope of Thermoproteus tenax: three-dimensional structure of the surface layer and its role in shape maintenance. *EMBO J* **6**, 1475-1480 (1987).

59 Danev, R. & Baumeister, W. Expanding the boundaries of cryo-EM with phase plates. *Curr Opin Struct Biol* **46**, 87-94, doi:10.1016/j.sbi.2017.06.006 (2017).

60 Castelle, C. J. & Banfield, J. F. Major New Microbial Groups Expand Diversity and Alter our Understanding of the Tree of Life. *Cell* **172**, 1181-1197, doi:10.1016/j.cell.2018.02.016 (2018).

61 Ponomarova, O. & Patil, K. R. Metabolic interactions in microbial communities: untangling the Gordian knot. *Curr Opin Microbiol* **27**, 37-44, doi:10.1016/j.mib.2015.06.014 (2015).

62 Mastronarde, D. N. SerialEM: a program for automated tilt series acquisition on Tecnai microscopes using prediction of specimen position. *Microscopy and Microanalysis* **9**, 1182-1183 (2003).

63 Kremer, J. R., Mastronarde, D. N. & McIntosh, J. R. Computer visualization of three-dimensional image data using IMOD. *J Struct Biol* **116**, 71-76, doi:10.1006/jsbi.1996.0013 (1996).

64 Tang, G. *et al.* EMAN2: an extensible image processing suite for electron microscopy. *J Struct Biol* **157**, 38-46, doi:10.1016/j.jsb.2006.05.009 (2007).

65 Galaz-Montoya, J. G., Flanagan, J., Schmid, M. F. & Ludtke, S. J. Single particle tomography in EMAN2. *J Struct Biol* **190**, 279-290, doi:10.1016/j.jsb.2015.04.016 (2015).

66 Galaz-Montoya, J. G. *et al.* Alignment algorithms and per-particle CTF correction for single particle cryo-electron tomography. *J Struct Biol* **194**, 383-394, doi:10.1016/j.jsb.2016.03.018 (2016).

67 Chen, M. *et al.* Convolutional neural networks for automated annotation of cellular cryo-electron tomograms. *Nat Methods* **14**, 983-985, doi:10.1038/nmeth.4405 (2017).

68 Pettersen, E. F. *et al.* UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605-1612, doi:10.1002/jcc.20084 (2004).

69 Danita, C., Chiu, W. & Galaz-Montoya, J. G. Efficient manual annotation of cryogenic electron tomograms using IMOD. *STAR Protoc* **3**, 101658, doi:10.1016/j.xpro.2022.101658 (2022).

70 Skinner, S. O., Sepulveda, L. A., Xu, H. & Golding, I. Measuring mRNA copy number in individual Escherichia coli cells using single-molecule fluorescent in situ hybridization. *Nat Protoc* **8**, 1100-1113, doi:10.1038/nprot.2013.066 (2013).

71 Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nat Methods* **9**, 676-682, doi:10.1038/nmeth.2019 (2012).

72 Yilmaz, P. *et al*. The SILVA and "All-species Living Tree Project (LTP)" taxonomic frameworks. *Nucl Acids Res* **42**, D643-D648 (2014).

73 McMurdie, P.J. & Holmes, S.P. phyloseq: An R Package for reproducible interactive analysis and graphics of microbiome census data. *PLoS ONE* **8(4),** e61217 (2013).

74 Woyke, T. *et al.* Decontamination of MDA reagents for single cell whole genome amplification. *PLoS One* **6**, e26161, doi:10.1371/journal.pone.0026161 (2011).

75 Bankevich, A. *et al.* SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* **19**, 455-477, doi:10.1089/cmb.2012.0021 (2012).

76 Hyatt, D. *et al.* Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* **11**, 119, doi:10.1186/1471-2105-11-119 (2010).

77 Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**, 357-359, doi:10.1038/nmeth.1923 (2012).

78 Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-2079, doi:10.1093/bioinformatics/btp352 (2009).

79 Eddy, S. R. Accelerated Profile HMM Searches. *PLoS Comput Biol* **7**, e1002195, doi:10.1371/journal.pcbi.1002195 (2011).

80 Dick, G. J. *et al.* Community-wide analysis of microbial genome sequence signatures. *Genome Biol* **10**, R85, doi:10.1186/gb-2009-10-8-r85 (2009).

81 Ultsch, A. & Mörchen, F. *ESOM-Maps: tools for clustering, visualization, and classification with Emergent SOM*. Vol. 46 (Univ., 2005).

82 Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P. & Tyson, G. W. CheckM: assessing the quality of microbial genomes recovered from isolates,

single cells, and metagenomes. *Genome Res* **25**, 1043-1055, doi:10.1101/gr.186072.114 (2015).

83 Hug, L. A. *et al.* Community genomic analyses constrain the distribution of metabolic traits across the Chloroflexi phylum and indicate roles in sediment carbon cycling. *Microbiome* **1**, 22, doi:10.1186/2049-2618-1-22 (2013).

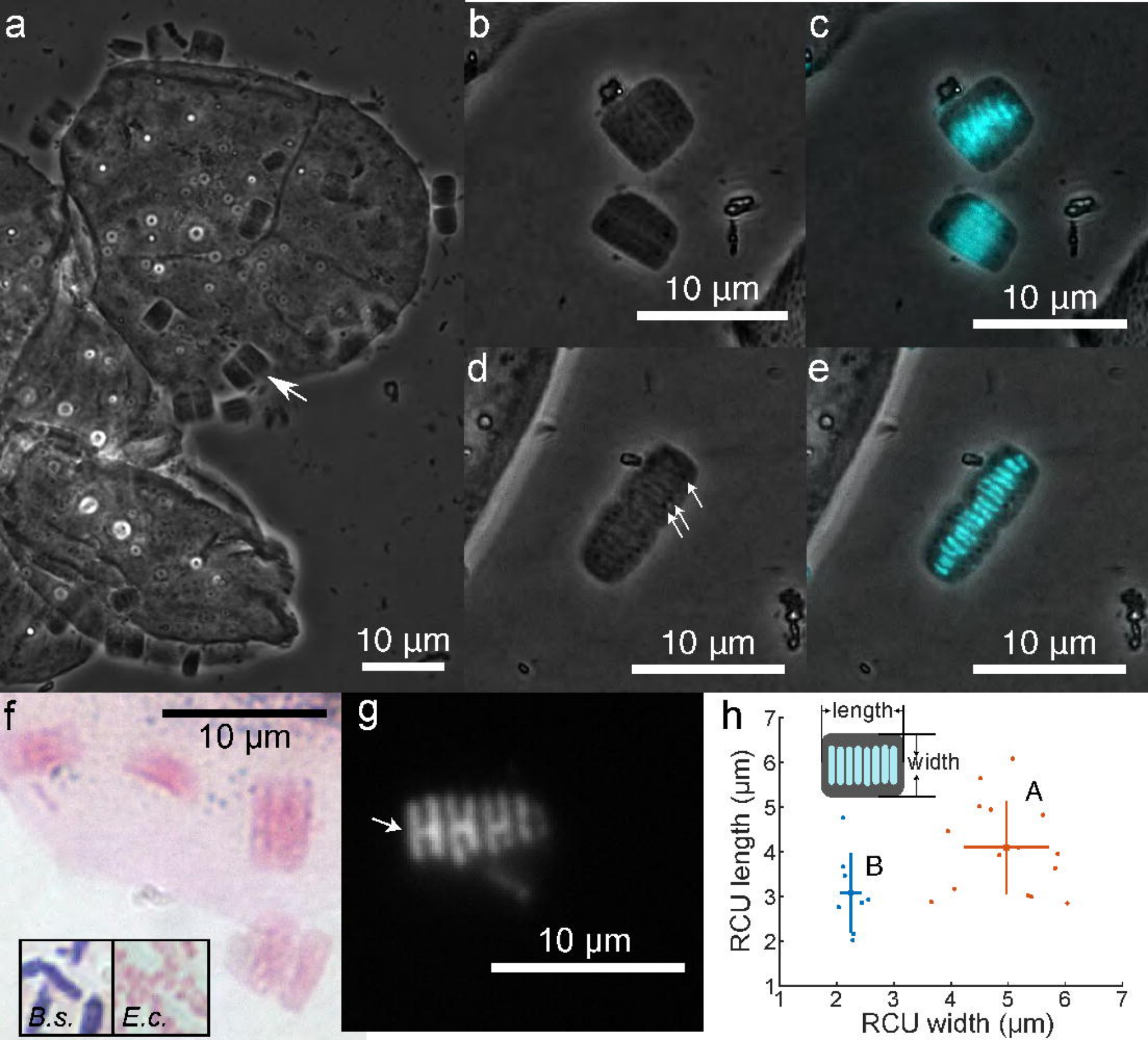
84 Finn, R. D. *et al.* The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res* **44**, D279-285, doi:10.1093/nar/gkv1344 (2016).

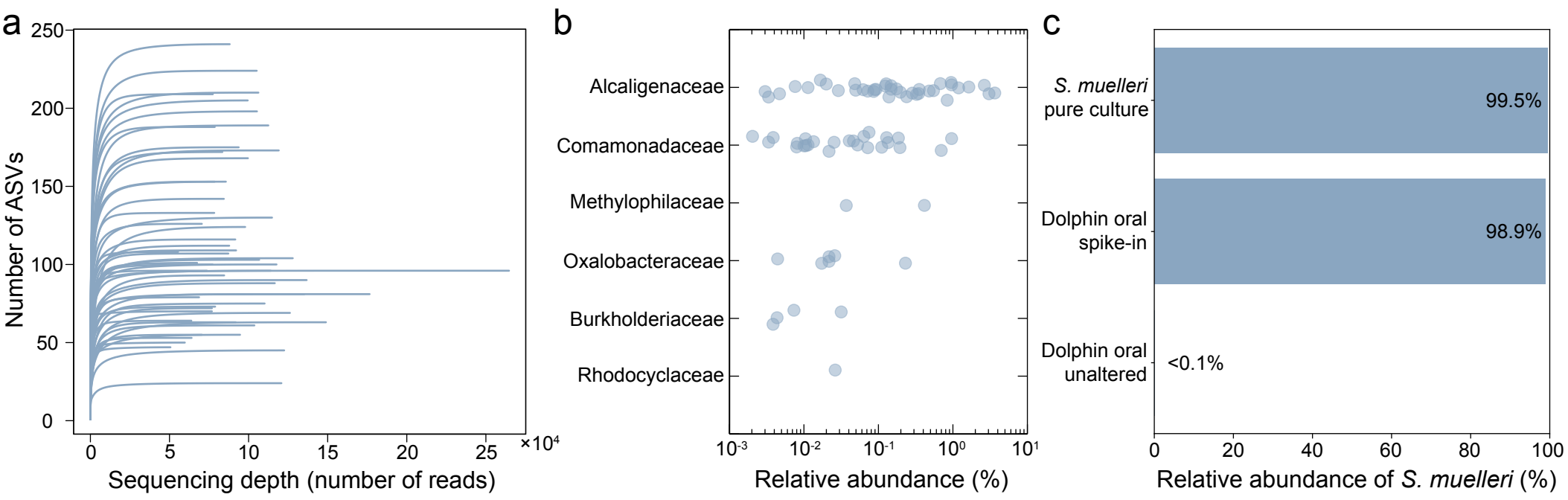
85 Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J Mol Biol* **215**, 403-410, doi:10.1016/S0022-2836(05)80360-2 (1990).

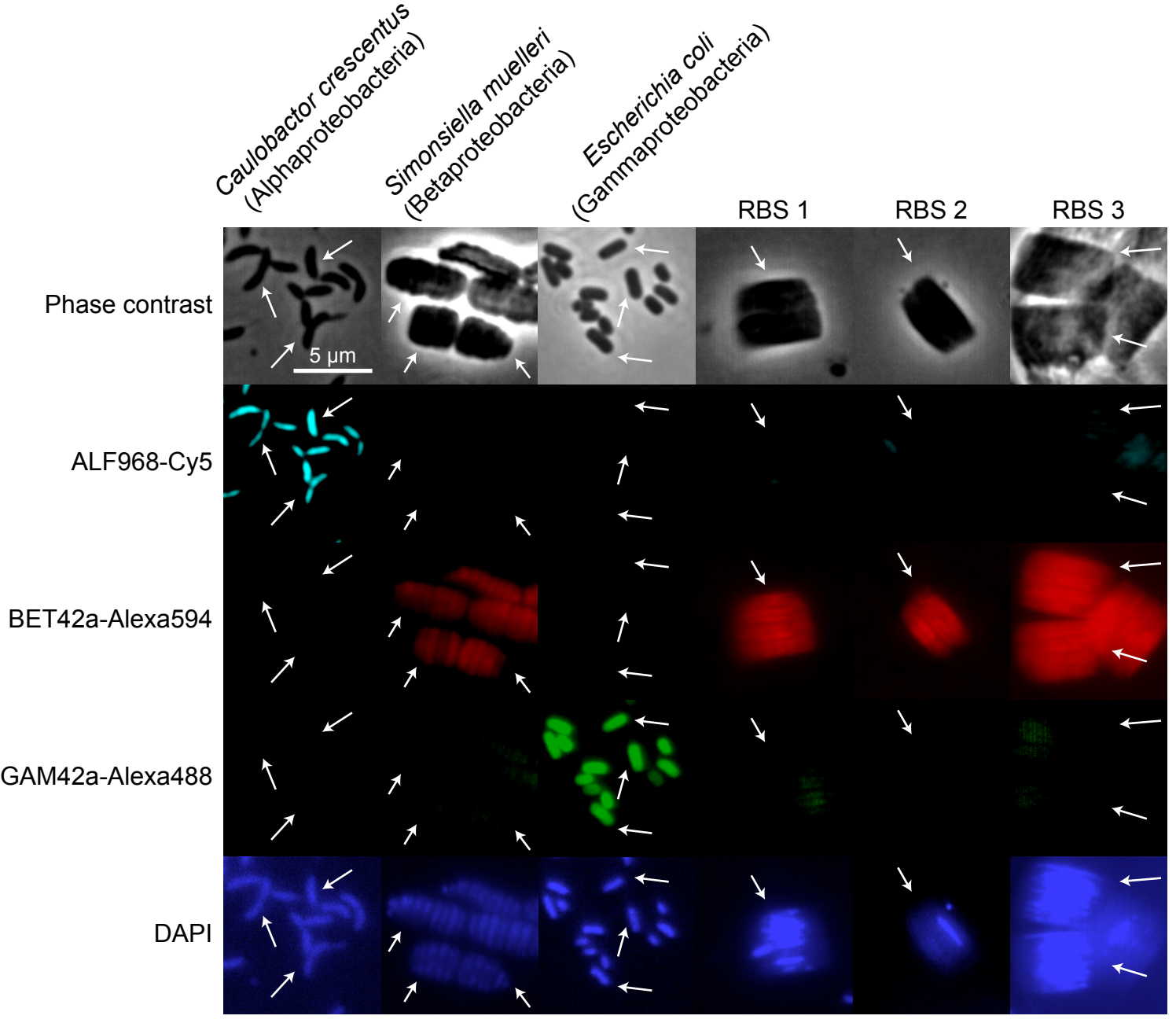
86 Kuhn, D. A., Gregory, D. A., Buchanan, G. E., Jr., Nyby, M. D. & Daly, K. R. Isolation, characterization, and numerical taxonomy of Simonsiella strains from the oral cavities of cats, dogs, sheep, and humans. *Arch Microbiol* **118**, 235-241, doi:10.1007/BF00429111 (1978).

87 Tian, Y. *et al.* Using DGGE profiling to develop a novel culture medium suitable for oral microbial communities. *Mol Oral Microbiol* **25**, 357-367, doi:10.1111/j.2041-1014.2010.00585.x (2010).

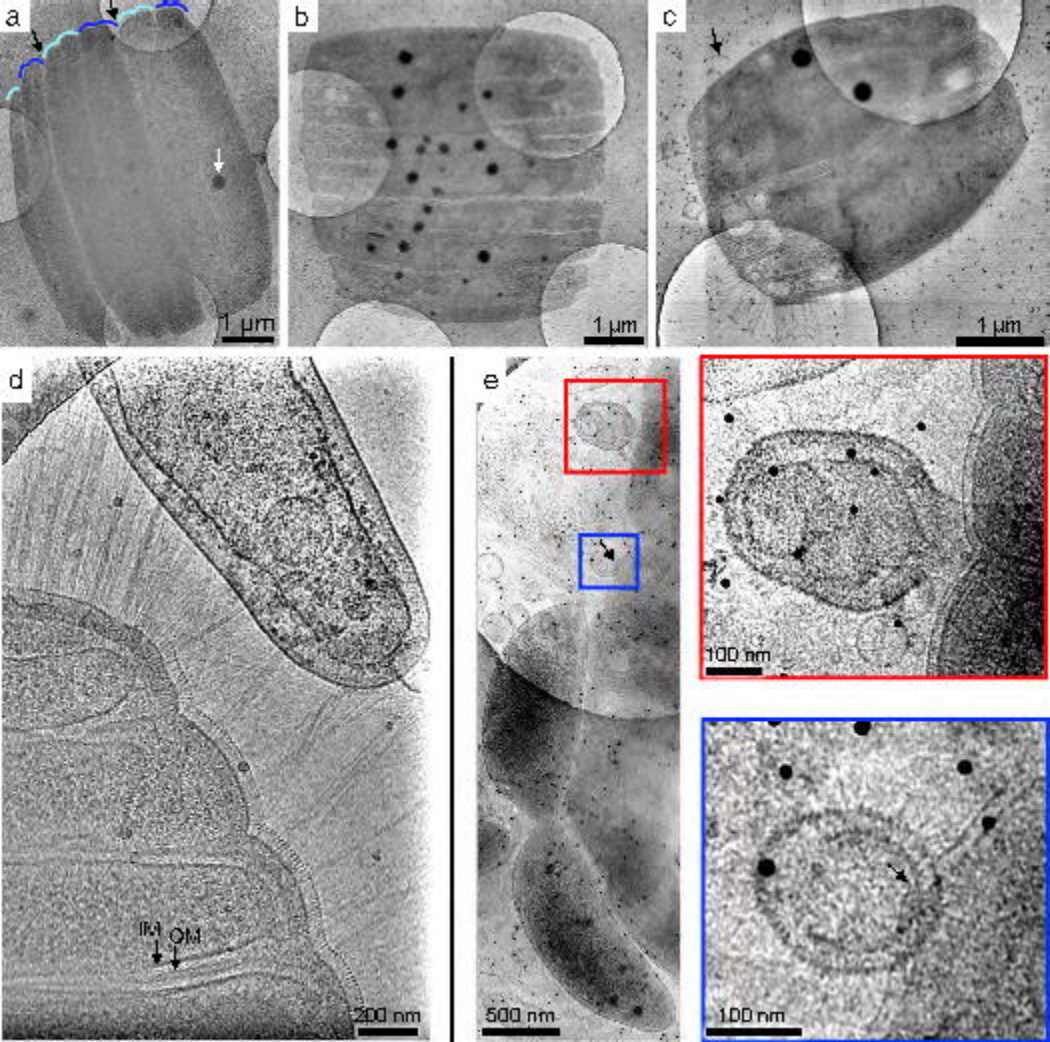
88 Meheust, R., Burstein, D., Castelle, C. J. & Banfield, J. F. The distinction of CPR bacteria from other bacteria based on protein family content. *Nat Commun* **10**, 4173, doi:10.1038/s41467-019-12171-z (2019).

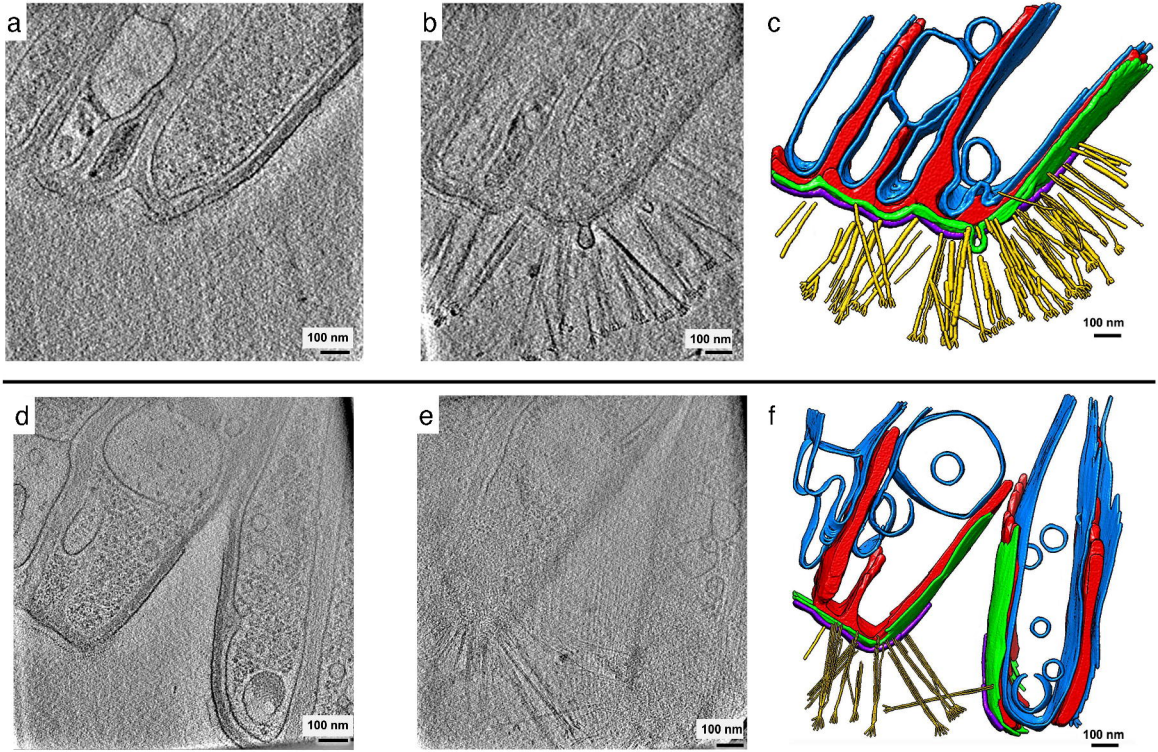

a
10 µm
b
10 µm
c
10 µm
d
10 µm
e
10 µm
f
10 µm
B.s.
E.c.
g
10 µm
h
length
width
A
B
RCU length (µm)
RCU width (µm)
1
2
3
4
5
6
7

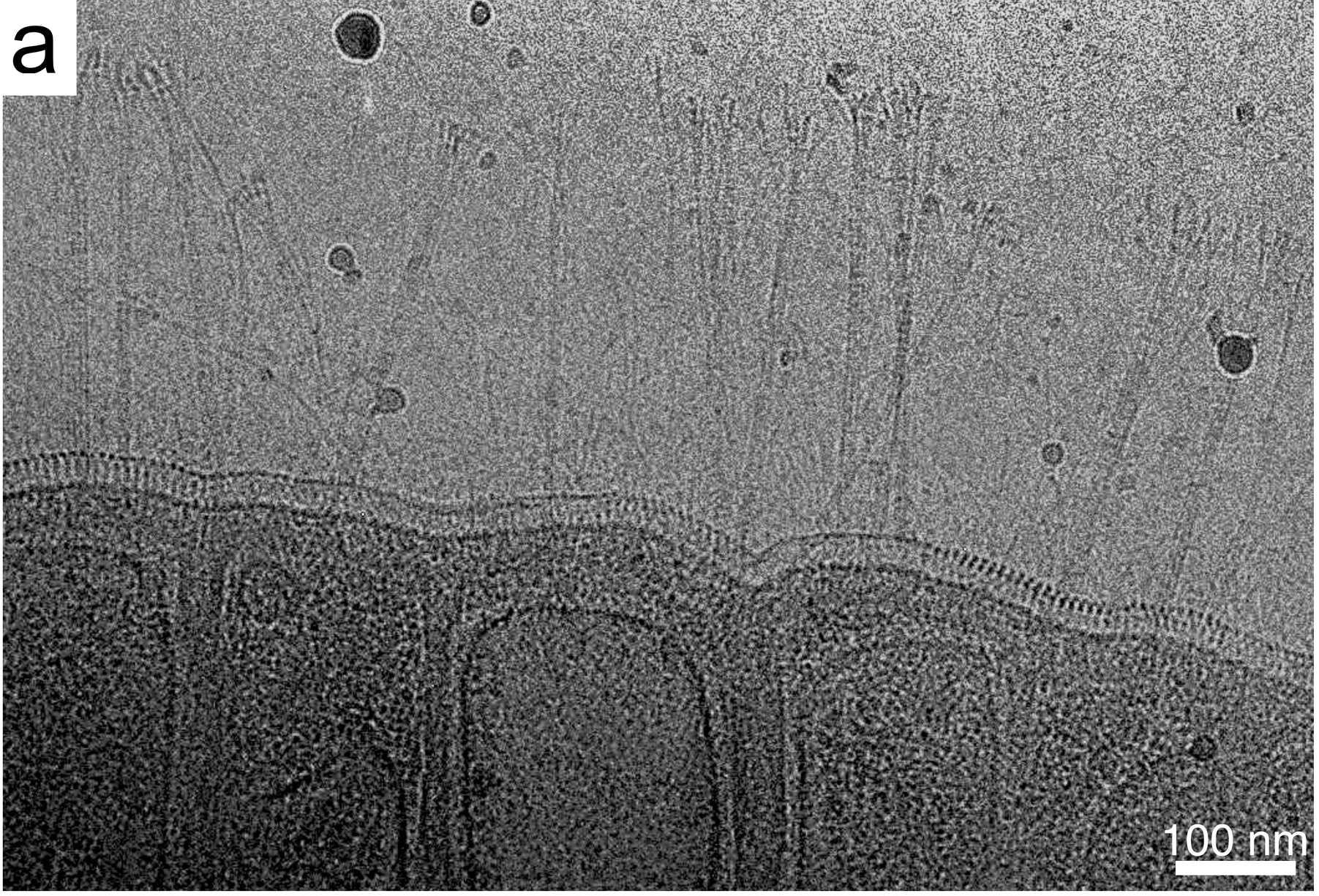

a
Number of ASVs
250
200
150
100
50
0
0
5
10
15
20
25
×10^4
Sequencing depth (number of reads)
b
Alcaligenaceae
Comamonadaceae
Methylophilaceae
Oxalobacteraceae
Burkholderiaceae
Rhodocyclaceae
10^-3
10^-2
10^-1
10^0
10^1
Relative abundance (%)
c
S. muelleri pure culture
99.5%
Dolphin oral spike-in
98.9%
Dolphin oral unaltered
<0.1%
0
20
40
60
80
100
Relative abundance of S. muelleri (%)

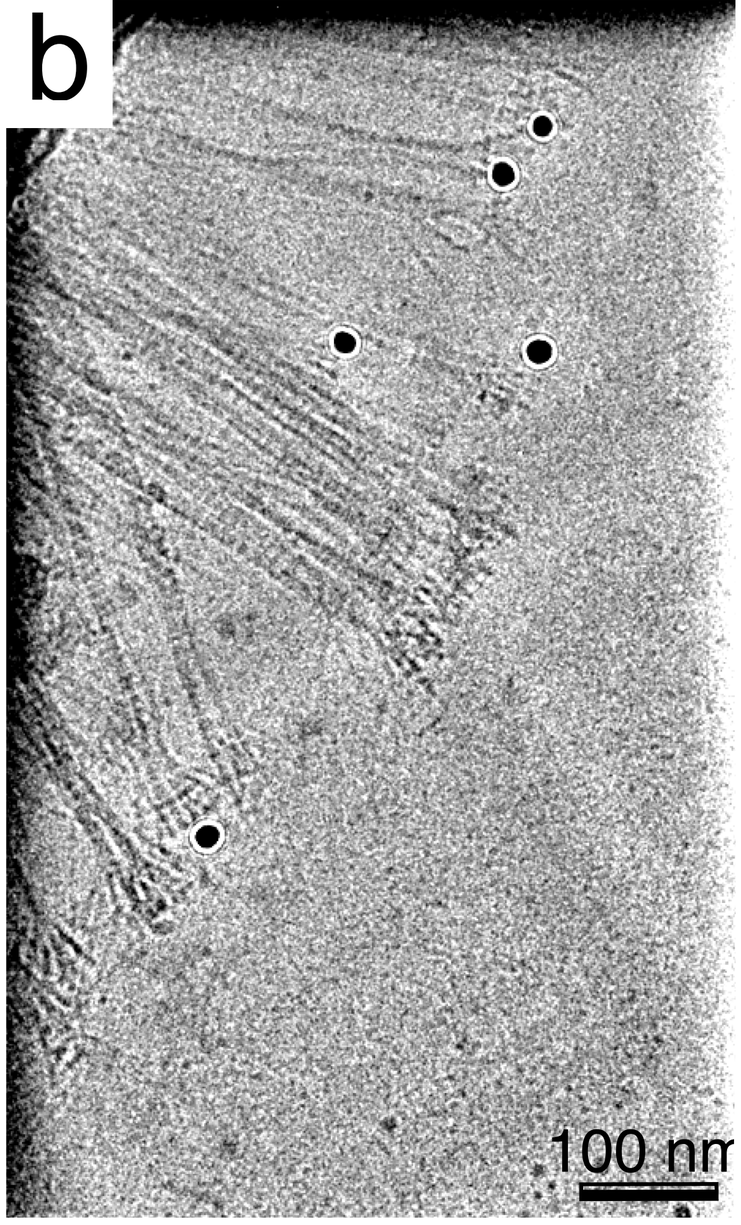

Caulobacter crescentus
(Alphaproteobacteria)
Simonsiella muelleri
(Betaproteobacteria)
Escherichia coli
(Gammaproteobacteria)
RBS 1
RBS 2
RBS 3
Phase contrast
5 µm
ALF968-Cy5
BET42a-Alexa594
GAM42a-Alexa488
DAPI

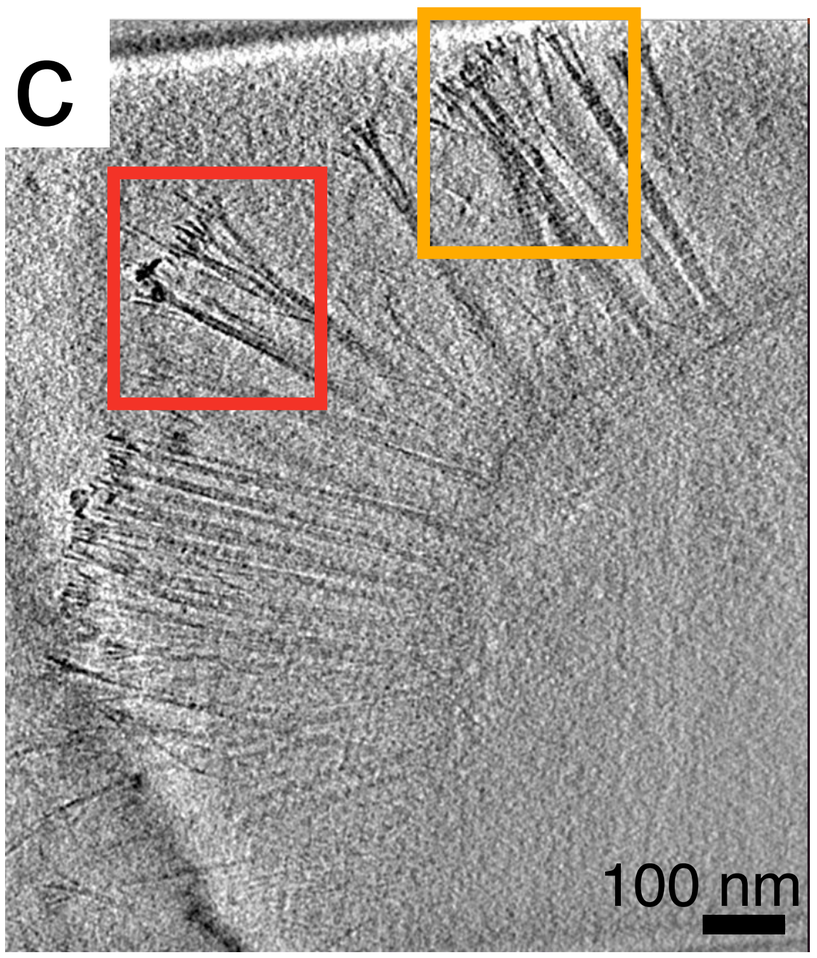
			RBS-A				Neg				Criteria			
Bin	**Phylum**	**Lowest taxonomic ID**	1	2	3	4	1	2	3	4	Genomics	Gram stain	16S 75%+	FISH
16	Betaproteobacteria	f_Alcaligenaceae											11	
4	Epsilonproteobacteria	g_Arcobacter											13	-
3	Epsilonproteobacteria	f_Campylobacteraceae*											13, 13	-
18	Gammaproteobacteria	p_Gammaproteobacteria											N/A	
11	Bacteroidetes	g_Tenacibaculum*											12, 12, 10	-
15	Gammaproteobacteria	p_Gammaproteobacteria											N/A	
13	Gammaproteobacteria	f_Moraxellaceae*											13, 13, 10	
17	Gammaproteobacteria	g_Pasteurella*											13, 13	
12	Bacteroidetes	f_Flavobacteriaceae											13	-
6	Fusobacteria	g_Fusobacterium											13	-
10	Gammaproteobacteria	p_Gammaproteobacteria											N/A	
9	Bacteroidetes	g_Porphyromonas*											11	-
14	Fusobacteria	g_Oceanivirga												-
5	Gracilibacteria	p_Gracilibacteria										†	10	-
2	Gracilibacteria	p_Gracilibacteria										†	10	-
7	Actinobacteria	f_Corynebacteriaceae*												-
8	Fungi	c_Malasseziomycetes											-	
1	Metazoa	s_Human											-	

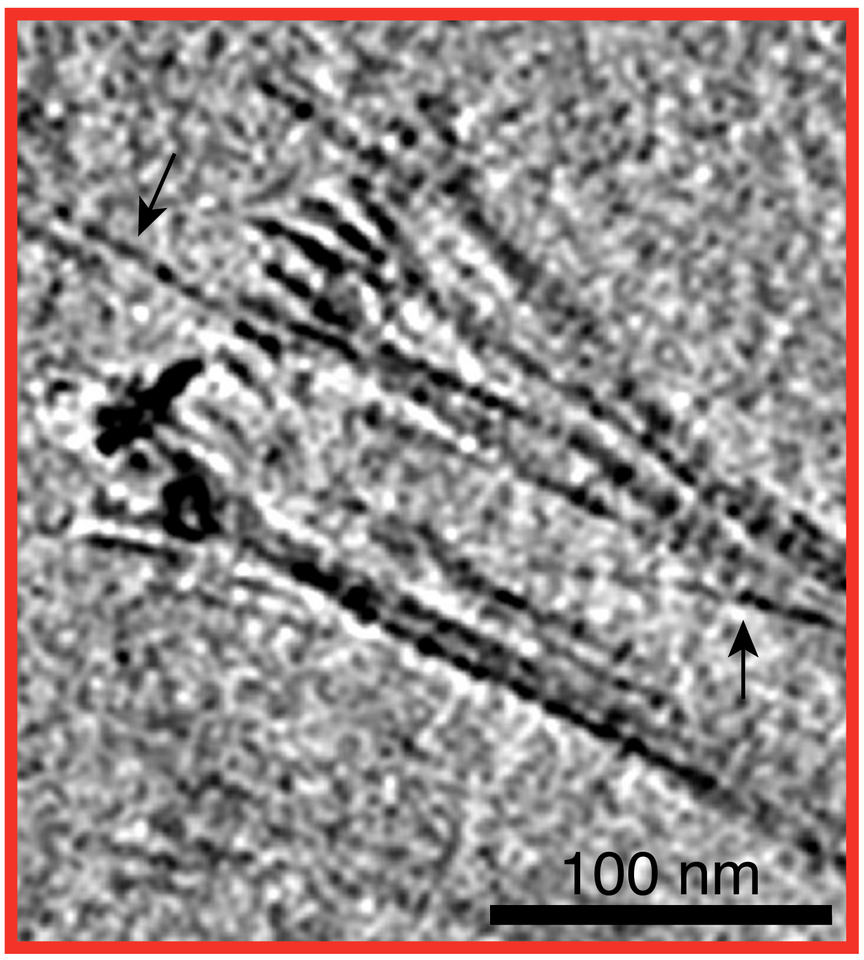

a
1 µm
b
1 µm
c
1 µm
d
IM
OM
200 nm
e
500 nm
100 nm
100 nm

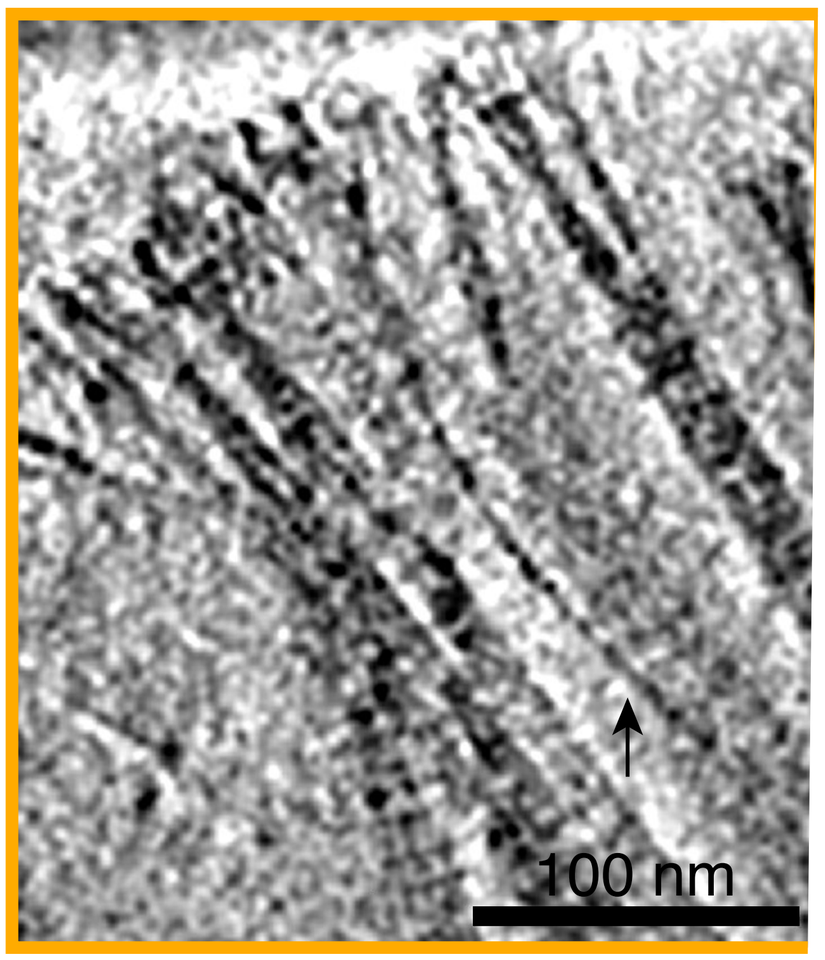

a
100 nm
b
100 nm
c
100 nm
d
100 nm
e
100 nm
f
100 nm

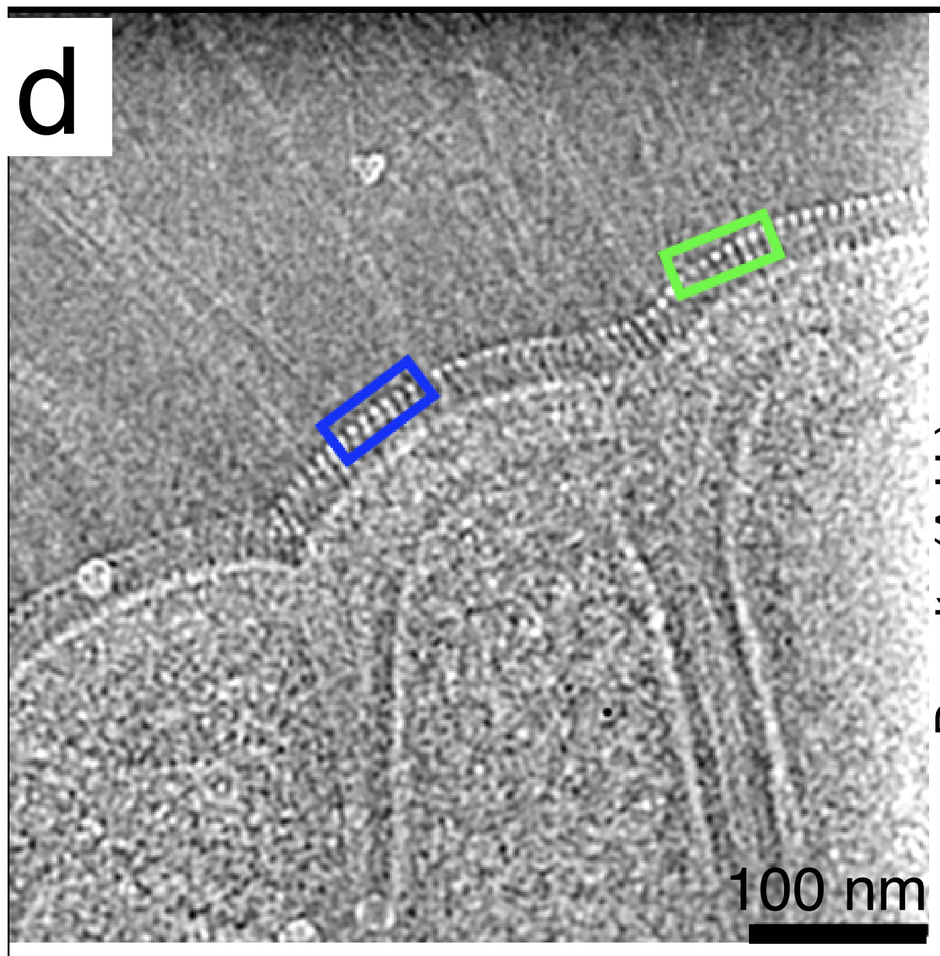

a
100 nm

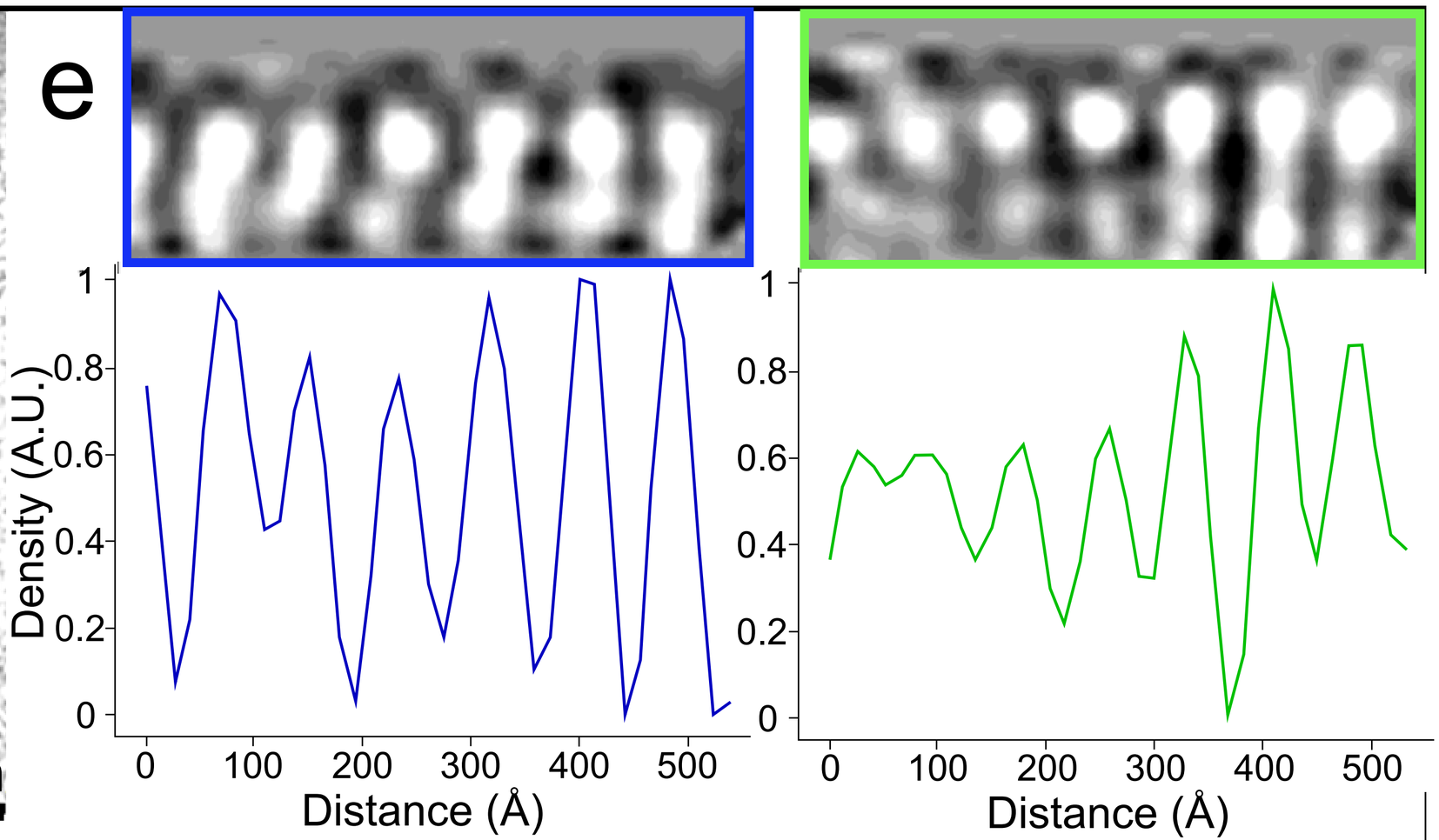

b
100 nm


c
100 nm


100 nm


100 nm


d
100 nm


e
Density (A.U.)
1
0.8
0.6
0.4
0.2
0
0
100
200
300
400
500
Distance (Å)
1
0.8
0.6
0.4
0.2
0
0
100
200
300
400
500
Distance (Å)